

FUNDACION MAS i MANJON
universidad de Stanford

Consciousness

First published Fri Jun 18, 2004; substantive revision Mon Aug 16, 2004

Perhaps no aspect of mind is more familiar or more puzzling than consciousness and our conscious experience of self and world. The problem of consciousness is arguably the central issue in current theorizing about the mind. Despite the lack of any agreed upon theory of consciousness, there is a widespread, if less than universal, consensus that an adequate account of mind requires a clear understanding of it and its place in nature. We need to understand both what consciousness is and how it relates to other, nonconscious, aspects of reality.

- 1. History of the issue
- 2. Concepts of consciousness
 - 2.1 Creature consciousness
 - 2.2 State consciousness
 - 2.3 Consciousness as an entity
- 3. Problems of consciousness
- 4. The descriptive question: What are the features of consciousness?
 - 4.1 First-person and third-person data
 - 4.2 Qualitative character
 - 4.3 Phenomenal structure
 - 4.4 Subjectivity
 - 4.5 Self-perspectival organization
 - 4.6 Unity
 - 4.7 Intentionality and transparency
 - 4.8 Dynamic flow
- 5. The explanatory question: How can consciousness exist?
 - 5.1 Diverse explanatory projects
 - 5.2 The explanatory gap
 - 5.3 Reductive and non-reductive explanation
 - 5.4 Prospects for explanatory success
- 6. The functional question: Why does consciousness exist?
 - 6.1 Causal status of consciousness
 - 6.2 Flexible control
 - 6.3 Social coordination
 - 6.4 Integrated representation
 - 6.5 Informational access
 - 6.6 Freedom of will

- 6.7 Intrinsic motivation
- 6.8 Constitutive and contingent roles
- 7. Theories of consciousness
- 8. Metaphysical theories of consciousness
 - 8.1 Dualist theories
 - 8.2 Physicalist theories
- 9. Specific theories of consciousness
 - 9.1 Higher-order theories
 - 9.2 Representational theories
 - 9.3 Cognitive theories
 - 9.4 Neural theories
 - 9.5 Quantum theories
 - 9.6 Nonphysical theories
- 10. Conclusion
- Bibliography
- Other Internet Resources
- Related Entries

1. History of the issue

Questions about the nature of conscious awareness have likely been asked for as long as there have been humans. Neolithic burial practices appear to express spiritual beliefs and provide early evidence for at least minimally reflective thought about the nature of human consciousness (Pearson 1999, Clark and Riel-Salvatore 2001). Preliterate cultures have similarly been found invariably to embrace some form of spiritual or at least animist view that indicates a degree of reflection about the nature of conscious awareness.

Nonetheless, some have argued that consciousness as we know it today is a relatively recent historical development that arose sometime after the Homeric era (Jaynes 1974). According to this view, earlier humans including those who fought the Trojan War did not experience themselves as unified internal subjects of their thoughts and actions, at least not in the ways we do today. Others have claimed that even during the classical period, there was no word of ancient Greek that corresponds to “consciousness” (Wilkes 1984, 1988, 1995). Though the ancients had much to say about mental matters, it is less clear whether they had any specific concepts or concerns for what we now think of as consciousness.

Although the words “conscious” and “conscience” are used quite differently today, it is likely that the Reformation emphasis on the latter as an inner source of truth played some role in the inward turn so characteristic of the modern reflective view of self. The Hamlet who walked the stage in 1600 already saw his world and self with profoundly modern

eyes.

By the beginning of the early modern era in the seventeenth century, consciousness had come full center in thinking about the mind. Indeed from the mid-17th through the late 19th century, consciousness was widely regarded as essential or definitive of the mental. René Descartes defined the very notion of thought (*pensée*) in terms of reflexive consciousness or self-awareness. In the *Principles of Philosophy* (1640) he wrote,

By the word 'thought' ('*pensée*') I understand all that of which we are conscious as operating in us.

Later, toward the end of the 17th century, John Locke offered a similar if slightly more qualified claim in *An Essay on Human Understanding* (1688),

I do not say there is no soul in man because he is not sensible of it in his sleep. But I do say he can not think at any time, waking or sleeping, without being sensible of it. Our being sensible of it is not necessary to anything but our thoughts, and to them it is and to them it always will be necessary.

Locke explicitly forswore making any hypothesis about the substantial basis of consciousness and its relation to matter, but he clearly regarded it as essential to thought as well as to personal identity.

Locke's contemporary G.W. Leibniz, drawing possible inspiration from his mathematical work on differentiation and integration, offered a theory of mind in the *Discourse on Metaphysics* (1686) that allowed for infinitely many degrees of consciousness and perhaps even for some thoughts that were unconscious, the so called "*petites perceptions*". Leibniz was the first to distinguish explicitly between perception and apperception, i.e., roughly between awareness and self-awareness. In the *Monadology* (1720) he also offered his famous analogy of the mill to express his belief that consciousness could not arise from mere matter. He asked his reader to imagine someone walking through an expanded brain as one would walk through a mill and observing all its mechanical operations, which for Leibniz exhausted its physical nature. Nowhere, he asserts, would such an observer see any conscious thoughts.

Despite Leibniz's recognition of the possibility of unconscious thought, for most of the next two centuries the domains of thought and consciousness were regarded as more or less the same. Associationist psychology, whether pursued by Locke or later in the eighteenth century by David Hume (1739) or in the nineteenth by James Mill (1829), aimed to discover the principles by which conscious thoughts or ideas interacted or affected each other. James Mill's son, John Stuart Mill continued his father's work on associationist psychology, but he allowed that combinations of ideas might produce resultants that went beyond their constituent mental parts, thus providing an early model of mental emergence (1865).

The purely associationist approach was critiqued in the late eighteenth century by Immanuel Kant (1787), who argued that an adequate account of experience and phenomenal consciousness required a far richer structure of mental and intentional organization. Phenomenal consciousness according to Kant could not be a mere succession of associated ideas, but at a minimum had to be the experience of a conscious self situated in an objective world structured with respect to space, time and causality.

Within the Anglo-American world, associationist approaches continued to be influential in both philosophy and psychology well into the twentieth century, while in the German and European sphere there was a greater interest in the larger structure of experience that led in part to the study of phenomenology through the work of Edmund Husserl (1913, 1929), Martin Heidegger (1927), Maurice Merleau-Ponty (1945) and others who expanded the study of consciousness into the realm of the social, the bodily and the interpersonal.

At the outset of modern scientific psychology in the mid-nineteenth century, the mind was still largely equated with consciousness, and introspective methods dominated the field as in the work of Wilhelm Wundt (1897), Hermann von Helmholtz (1897), William James (1890) and Alfred Titchener (1901). However, the relation of consciousness to brain remained very much a mystery as expressed in T. H. Huxley's famous remark,

How it is that anything so remarkable as a state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of the Djinn, when Aladdin rubbed his lamp (1866).

The early twentieth century saw the eclipse of consciousness from scientific psychology, especially in the United States with the rise of behaviorism (Watson 1924, Skinner 1953) though movements such as Gestalt psychology kept it a matter of ongoing scientific concern in Europe (Köhler 1929, Köffka 1935). In the 1960s, the grip of behaviorism weakened with the rise of cognitive psychology and its emphasis on information processing and the modeling of internal mental processes (Neisser 1965, Gardiner 1985). However, despite the renewed emphasis on explaining cognitive capacities such as memory, perception and language comprehension, consciousness remained a largely neglected topic for several further decades.

In the 1980s and 90s there was a major resurgence of scientific and philosophical research into the nature and basis of consciousness (Baars 1988, Dennett 1991, Penrose 1989, 1994, Crick 1994, Lycan 1987, 1996, Chalmers 1996). Once consciousness was back under discussion, there was a rapid proliferation of research with a flood of books and articles, as well as the introduction of specialty journals (*The Journal of Consciousness Studies*, *Consciousness and Cognition*, *Psyche*), professional societies (Association for the Scientific Study of Consciousness - ASSC) and annual conferences devoted exclusively to its investigation (*Toward a Science of Consciousness*, ASSC).

2. Concepts of Consciousness

The words “conscious” and “consciousness” are umbrella terms that cover a wide variety of mental phenomena. Both are used with a diversity of meanings, and the adjective “conscious” is heterogeneous in its range, being applied both to whole organisms — creature consciousness — and to particular mental states and processes — state consciousness (Rosenthal 1986, Gennaro 1995, Carruthers 2000).

2.1 Creature Consciousness

An animal, person or other cognitive system may be regarded as conscious in a number of different senses.

Sentience. It may be conscious in the generic sense of simply being a sentient creature, one capable of sensing and responding to its world (Armstrong 1981). Being conscious in this sense may admit of degrees, and just what sort of sensory capacities are sufficient may not be sharply defined. Are fish conscious in the relevant respect? And what of shrimp or bees?

Wakefulness. One might further require that the organism actually be exercising such a capacity rather than merely having the ability or disposition to do so. Thus one might count it as conscious only if it were awake and normally alert. In that sense organisms would not count as conscious when asleep or in any of the deeper levels of coma. Again boundaries may be blurry, and intermediate cases may be involved. For example, is one conscious in the relevant sense when dreaming, hypnotized or in a fugue state?

Self-consciousness. A third and yet more demanding sense might define conscious creatures as those that are not only aware but also aware that they are aware, thus treating creature consciousness as a form of self-consciousness (Carruthers 2000). The self-awareness requirement might get interpreted in a variety of ways, and which creatures would qualify as conscious in the relevant sense will vary accordingly. If it is taken to involve explicit conceptual self-awareness, many non-human animals and even young children might fail to qualify, but if only more rudimentary implicit forms of self-awareness are required then a wide range of nonlinguistic creatures might count as self-conscious.

What it is like. Thomas Nagel's (1974) famous “what it is like” criterion aims to capture another and perhaps more subjective notion of being a conscious organism. According to Nagel, a being is conscious just if there is “something that it is like” to be that creature, i.e., some subjective way the world seems or appears from the creature's mental or experiential point of view. In Nagel's example, bats are conscious because there is something that it is like for a bat to experience its world through its echo-locatory senses, even though we humans from our human point of view can not emphatically understand what such a mode of consciousness is like from the bat's own point of view.

Subject of conscious states. A fifth alternative would be to define the notion of a conscious organism in terms of conscious states. That is, one might first define what makes a mental state a conscious mental state, and then define being a conscious creature in terms of having such states. One's concept of a conscious organism would then depend upon the particular account one gives of conscious states (section 2.2).

Transitive Consciousness. In addition to describing creatures as conscious in these various senses, there are also related senses in which creatures are described as being conscious of various things. The distinction is sometimes marked as that between transitive and intransitive notions of consciousness, with the former involving some object at which consciousness is directed (Rosenthal 1986).

2.2 State consciousness.

The notion of a conscious mental state also has a variety of distinct though perhaps interrelated meanings. There are at least six major options.

States one is aware of. On one common reading, a conscious mental state is simply a mental state one is aware of being in (Rosenthal 1986, 1996). Conscious states in this sense involve a form of meta-mentality or meta-intentionality in so far as they require mental states that are themselves about mental states. To have a conscious desire for a cup of coffee is to have such a desire and also to be simultaneously and directly aware that one has such a desire. Unconscious thoughts and desires in this sense are simply those we have without being aware of having them, whether our lack of self-knowledge results from simple inattention or more deeply psychoanalytic causes.

Qualitative states. States might also be regarded as conscious in a seemingly quite different and more qualitative sense. That is, one might count a state as conscious just if it has or involves qualitative or experiential properties of the sort often referred to as "qualia" or "raw sensory feels". (See the entry on qualia.) One's perception of the Merlot one is drinking or of the fabric one is examining counts as a conscious mental state in this sense because it involves various sensory qualia, e.g., taste qualia in the wine case and color qualia in one's visual experience of the cloth. There is considerable disagreement about the nature of such qualia (Churchland 1985, Shoemaker 1990, Clark 1993, Chalmers 1996) and even about their existence. Traditionally qualia have been regarded as intrinsic, private, ineffable monadic features of experience, but current theories of qualia often reject at least some of those commitments (Dennett 1990).

Phenomenal states. Such qualia are sometimes referred to as phenomenal properties and the associated sort of consciousness as phenomenal consciousness, but the latter term is perhaps more properly applied to the overall structure of experience and involves far more than sensory qualia. The phenomenal structure of consciousness also encompasses much of the spatial, temporal and conceptual organization of our experience of the world

and of ourselves as agents in it. (See section 4.3) It is therefore probably best, at least initially, to distinguish the concept of phenomenal consciousness from that of qualitative consciousness, though they no doubt overlap.

What-it-is-like states. Consciousness in both those senses links up as well with Thomas Nagel's (1974) notion of a conscious creature, insofar as one might count a mental state as conscious in the "what it is like" sense just if there is something that it is like to be in that state. Nagel's criterion might be understood as aiming to provide a first-person or internal conception of what makes a state a phenomenal or qualitative state.

Access consciousness. States might be conscious in a seemingly quite different access sense, which has more to do with intra-mental relations. In this respect, a state's being conscious is a matter of its availability to interact with other states and of the access that one has to its content. In this more functional sense, which corresponds to what Ned Block (1995) calls access consciousness, a visual state's being conscious is not so much a matter of whether or not it has a qualitative "what it's likeness", but of whether or not it and the visual information that it carries is generally available for use and guidance by the organism. In so far as the information in that state is richly and flexibly available to its containing organism, then it counts as a conscious state in the relevant respect, whether or not it has any qualitative or phenomenal feel in the Nagel sense.

Narrative consciousness. States might also be regarded as conscious in a narrative sense that appeals to the notion of the "stream of consciousness", regarded as an ongoing more or less serial narrative of episodes from the perspective of an actual or merely virtual self. The idea would be to equate the person's conscious mental states with those that appear in the stream (Dennett 1991, 1992).

Although these six notions of what makes a state conscious can be independently specified, they are obviously not without potential links, nor do they exhaust the realm of possible options. Drawing connections, one might argue that states appear in the stream of consciousness only in so far as we are aware of them, and thus forge a bond between the first meta-mental notion of a conscious state and the stream or narrative concept. Or one might connect the access with the qualitative or phenomenal notions of a conscious state by trying to show that states that represent in those ways make their contents widely available in the respect required by the access notion.

Aiming to go beyond the six options, one might distinguish conscious from nonconscious states by appeal to aspects of their intra-mental dynamics and interactions other than mere access relations; e.g., conscious states might manifest a richer stock of content-sensitive interactions or a greater degree of flexible purposive guidance of the sort associated with the self-conscious control of thought. Alternatively, one might try to define conscious states in terms of conscious creatures. That is, one might give some account of what it is to be a conscious creature or perhaps even a conscious self, and then define one's notion of a conscious state in terms of being a state of such a creature or system, which would be the

converse of the last option considered above for defining conscious creatures in terms of conscious mental states.

2.3 Consciousness as an entity.

The noun “consciousness” has an equally diverse range of meanings that largely parallel those of the adjective “conscious”. Distinctions can be drawn between creature and state consciousness as well as among the varieties of each. One can refer specifically to phenomenal consciousness, access consciousness, reflexive or meta-mental consciousness, and narrative consciousness among other varieties.

Here consciousness itself is not typically treated as a substantive entity but merely the abstract reification of whatever property or aspect is attributed by the relevant use of the adjective “conscious”. Access consciousness is just the property of having the required sort of internal access relations, and qualitative consciousness is simply the property that is attributed when “conscious” is applied in the qualitative sense to mental states. How much this commits one to the ontological status of consciousness per se will depend on how much of a Platonist one is about universals in general. (See the entry on the medieval problem of universals.) It need not commit one to consciousness as a distinct entity any more than one's use of “square”, “red” or “gentle” commits one to the existence of squareness, redness or gentleness as distinct entities.

Though it is not the norm, one could nonetheless take a more robustly realist view of consciousness as a component of reality. That is one could think of consciousness as more on a par with electromagnetic fields than with life.

Since the demise of vitalism, we do not think of life per se as something distinct from living things. There are living things including organisms, states, properties and parts of organisms, communities and evolutionary lineages of organisms, but life is not itself a further thing, an additional component of reality, some vital force that gets added into living things. We apply the adjectives “living” and “alive” correctly to many things, and in doing so we might be said to be attributing life to them but with no meaning or reality other than that involved in their being living things.

Electromagnetic fields by contrast are regarded as real and independent parts of our physical world. Even though one may sometimes be able to specify the values of such a field by appeal to the behavior of particles in it, the fields themselves are regarded as concrete constituents of reality and not merely as abstractions or sets of relations among particles.

Similarly one could regard “consciousness” as referring to a component or aspect of reality that manifests itself in conscious states and creatures but is more than merely the abstract nominalization of the adjective “conscious” we apply to them. Though such strongly realist views are not very common at present, they should be included within the logical

space of options.

There are thus many concepts of consciousness, and both “conscious” and “consciousness” are used in a wide range of ways with no privileged or canonical meaning. However, this may be less of an embarrassment than an embarrassment of riches. Consciousness is a complex feature of the world, and understanding it will require a diversity of conceptual tools for dealing with its many differing aspects. Conceptual plurality is thus just what one would hope for. As long as one avoids confusion by being clear about one's meanings, there is great value in having a variety of concepts by which we can access and grasp consciousness in all its rich complexity. However, one should not assume that conceptual plurality implies referential divergence. Our multiple concepts of consciousness may in fact pick out varying aspects of a single unified underlying mental phenomenon. Whether and to what extent they do so remains an open question.

3. Problems of Consciousness

The task of understanding consciousness is an equally diverse project. Not only do many different aspects of mind count as conscious in some sense, each is also open to various respects in which it might be explained or modeled. Understanding consciousness involves a multiplicity not only of explananda but also of questions that they pose and the sorts of answers they require. At the risk of oversimplifying, the relevant questions can be gathered under three crude rubrics as the What, How, and Why questions:

- The Descriptive Question: What is consciousness? What are its principal features? And by what means can they be best discovered, described and modeled?
- The Explanatory Question: How does consciousness of the relevant sort come to exist? Is it a primitive aspect of reality, and if not how does (or could) consciousness in the relevant respect arise from or be caused by nonconscious entities or processes?
- The Functional Question: Why does consciousness of the relevant sort exist? Does it have a function, and if so what it is it? Does it act causally and if so with sorts of effects? Does it make a difference to the operation of systems in which it is present, and if so why and how?

The three questions focus respectively on describing the features of consciousness, explaining its underlying basis or cause, and explicating its role or value. The divisions among the three are of course somewhat artificial, and in practice the answers one gives to each will depend in part on what one says about the others. One can not, for example, adequately answer the what question and describe the main features of consciousness without addressing the why issue of its functional role within systems whose operations it affects. Nor could one explain how the relevant sort of consciousness might arise from nonconscious processes unless one had a clear account of just what features had to be caused or realized to count as producing it. Those caveats notwithstanding, the three-way division of questions provides a useful structure for articulating the overall explanatory project and for assessing the adequacy of particular theories or models of consciousness.

4. The descriptive question: What are the features of consciousness?

The What question asks us to describe and model the principal features of consciousness, but just which features are relevant will vary with the sort of consciousness we aim to capture. The main properties of access consciousness may be quite unlike those of qualitative or phenomenal consciousness, and those of reflexive consciousness or narrative consciousness may differ from both. However, by building up detailed theories of each type, we may hope to find important links between them and perhaps even to discover that they coincide in at least some key respects.

4.1 First-person and third-person data

The general descriptive project will require a variety of investigational methods (Flanagan 1992). Though one might naively regard the facts of consciousness as too self-evident to require any systematic methods of gathering data, the epistemic task is in reality far from trivial (Husserl 1913).

First-person introspective access provides a rich and essential source of insight into our conscious mental life, but it is neither sufficient in itself nor even especially helpful unless used in a trained and disciplined way. Gathering the needed evidence about the structure of experience requires us both to become phenomenologically sophisticated self-observers and to complement our introspective results with many types of third-person data available to external observer (Searle 1992, Varela 1995, Siewert 1998)

As phenomenologists have known for more than a century, discovering the structure of conscious experience demands a rigorous inner-directed stance that is quite unlike our everyday form or self-awareness (Husserl 1929, Merleau-Ponty 1945). Skilled observation of the needed sort requires training, effort and the ability to adopt alternative perspectives on one's experience.

The need for third-person empirical data gathered by external observers is perhaps most obvious with regard to the more clearly functional types of consciousness such as access consciousness, but it is required even with regard to phenomenal and qualitative consciousness. For example, deficit studies that correlate various neural and functional sites of damage with abnormalities of conscious experience can make us aware of aspects of phenomenal structure that escape our normal introspective awareness. As such case studies show, things can come apart in experience that seem inseparably unified or singular from our normal first-person point of view (Sacks 1985, Shallice 1988, Farah 1995).

Or to pick another example, third-person data can make us aware of how our experiences of acting and our experiences of event-timing affect each other in ways that we could never discern through mere introspection (Libet 1985, Wegner 2002). Nor are the facts gathered by these third person methods merely about the causes or bases of consciousness; they often concern the very structure of phenomenal consciousness itself.

First-person, third-person and perhaps even second-person (Varela 1995) interactive methods will all be needed to collect the requisite evidence.

Using all these sources of data, we will hopefully be able to construct detailed descriptive models of the various sorts of consciousness. Though the specific features of most importance may vary among the different types, our overall descriptive project will need to address at least the following seven general aspects of consciousness (sections 4.2 - 4.7).

4.2 Qualitative character

Qualitative character is often equated with so called “raw feels” and illustrated by the redness one experiences when one looks at ripe tomatoes or the specific sweet sapor one encounters when one tastes an equally ripe pineapple (Locke 1688). The relevant sort of qualitative character is not restricted to sensory states, but is typically taken to be present as an aspect of experiential states in general, such as experienced thoughts or desires (Siewert 1998).

The existence of such feels may seem to some to mark the threshold for states or creatures that are really conscious. If an organism senses and responds in apt ways to its world but lacks such qualia, then it might count as conscious at best in a loose and less than literal sense. Or so at least it would seem to those who take qualitative consciousness in the “what it is like” sense to be philosophically and scientifically central (Nagel 1974, Chalmers 1996).

Qualia problems in many forms — Can there be inverted qualia? (Block 1980a 1980b, Shoemaker 1981, 1982) Are qualia epiphenomenal? (Jackson 1982, Chalmers 1996) How could neural states give rise to qualia? (Levine 1983, McGinn 1991) — have loomed large in the recent past. But the What question raises a more basic problem of qualia: namely that of giving a clear and articulated description of our qualia space and the status of specific qualia within it.

Absent such a model, factual or descriptive errors are all too likely. For example, claims about the unintelligibility of the link between experienced red and any possible neural substrate of such an experience sometimes treat the relevant color quale as a simple and sui generis property (Levine 1983), but phenomenal redness in fact exists within a complex color space with multiple systematic dimensions and similarity relations (Hardin 1992). Understanding the specific color quale relative to that larger relational structure not only gives us a better descriptive grasp of its qualitative nature, it may also provide some “hooks” to which one might attach intelligible psycho-physical links.

Color may be the exception in terms of our having a specific and well developed formal understanding of the relevant qualitative space, but it is not likely an exception with regard to the importance of such spaces to our understanding of qualitative properties in general (Clark 1993, P.M. Churchland 1995). (See the entry on qualia.)

4.3 Phenomenal structure

Phenomenal structure should not be conflated with qualitative structure, despite the sometimes interchangeable use of “qualia” and “phenomenal properties” in the literature. “Phenomenal organization” covers all the various kinds of order and structure found within the domain of experience, i.e., within the domain of the world as it appears to us. There are obviously important links between the phenomenal and the qualitative. Indeed qualia might be best understood as properties of phenomenal or experienced objects, but there is in fact far more to the phenomenal than raw feels. As Kant (1787), Husserl (1913), and generations of phenomenologists have shown, the phenomenal structure of experience is richly intentional and involves not only sensory ideas and qualities but complex representations of time, space, cause, body, self, world and the organized structure of lived reality in all its conceptual and nonconceptual forms.

Since many non-conscious states also have intentional and representational aspects, it may be best to consider phenomenal structure as involving a special kind of intentional and representational organization and content, the kind distinctively associated with consciousness (Siewert 1998). (See the entry on representational theories of consciousness).

Answering the What question requires a careful account of the coherent and densely organized representational framework within which particular experiences are embedded. Since most of that structure is only implicit in the organization of experience, it can not just be read off by introspection. Articulating the structure of the phenomenal domain in a clear and intelligible way is a long and difficult process of inference and model building (Husserl 1929). Introspection can aid it, but a lot of theory construction and ingenuity are also needed.

4.4 Subjectivity

Subjectivity is another notion sometimes equated with the qualitative or the phenomenal aspects of consciousness in the literature, but again there are good reason to recognize it, at least in some of its forms, as a distinct feature of consciousness — related to the qualitative and the phenomenal but different from each. In particular, the epistemic form of subjectivity concerns apparent limits on the knowability or even the understandability of various facts about conscious experience (Nagel 1974, Van Gulick 1985, Lycan 1996).

On Thomas Nagel's (1974) account, facts about what it is like to be a bat are subjective in the relevant sense because they can be fully understood only from the bat-type point of view. Only creatures capable of having or undergoing similar such experiences can understand their what-it's-likeness in the requisite empathetic sense. Facts about conscious experience can be best incompletely understood from an outside third person point of view, such as those associated with objective physical science. A similar view about the limits of third-person theory seems to lie behind claims regarding what Frank Jackson's

(1982) hypothetical Mary, the super color scientist, could not understand about experiencing red because of her own impoverished history of achromatic visual experience.

Whether facts about experience are indeed epistemically limited in this way is open to debate (Lycan 1996), but the claim that understanding consciousness requires special forms of knowing and access from the inside point of view is intuitively plausible and has a long history (Locke 1688). Thus any adequate answer to the What question must address the epistemic status of consciousness, both our abilities to understand it and their limits (Papineau 2002, Chalmers 2003). (See the entry on self-knowledge).

4.5 Self-perspectival organization

The perspectival structure of consciousness is one aspect of its overall phenomenal organization, but it is important enough to merit discussion in its own right. Insofar as the key perspective is that of the conscious self, the specific feature might be called self-perspectuality. Conscious experiences do not exist as isolated mental atoms, but as modes or states of a conscious self or subject (Descartes 1644, Searle 1992, though pace Hume 1739). A visual experience of a blue sphere is always a matter of there being some self or subject who is appeared to in that way. A sharp and stabbing pain is always a pain felt or experienced by some conscious subject. The self need not appear as an explicit element in our experiences, but as Kant (1787) noted the “I think” must at least potentially accompany each of them.

The self might be taken as the perspectival point from which the world of objects is present to experience (Wittgenstein 1921). It provides not only a spatial and temporal perspective for our experience of the world but one of meaning and intelligibility as well. The intentional coherence of the experiential domain relies upon the dual interdependence between self and world: the self as perspective from which objects are known and the world as the integrated structure of objects and events whose possibilities of being experienced implicitly define the nature and location of the self (Kant 1787, Husserl 1929).

Conscious organisms obviously differ in the extent to which they constitute a unified and coherent self, and they likely differ accordingly in the sort or degree of perspectival focus they embody in their respective forms of experience (Lorenz 1977). Consciousness may not require a distinct or substantial self of the traditional Cartesian sort, but at least some degree of perspectivally self-like organization seems essential for the existence of anything that might count as conscious experience. Experiences seem no more able to exist without a self or subject to undergo them than could ocean waves exist without the sea through which they move. The Descriptive question thus requires some account of the self-perspectival aspect of experience and the self-like organization of conscious minds on which it depends, even if the relevant account treats the self in a relatively deflationary and virtual way (Dennett 1991, 1992).

4.6 Unity

Unity is closely linked with the self-perspective, but it merits specific mention on its own as a key aspect of the organization of consciousness. Conscious systems and conscious mental states both involve many diverse forms of unity. Some are causal unities associated with the integration of action and control into a unified focus of agency. Others are more representational and intentional forms of unity involving the integration of diverse items of content at many scales and levels of binding (Cleeremans 2003).

Some such integrations are relatively local as when diverse features detected within a single sense modality are combined into a representation of external objects bearing those features, e.g. when one has a conscious visual experience of a moving red soup can passing above a green striped napkin (Triesman and Gelade 1980).

Other forms of intentional unity encompass a far wider range of contents. The content of one's present experience of the room in which one sits depends in part upon its location within a far larger structure associated with one's awareness of one's existence as an ongoing temporally extended observer within a world of spatially connected independently existing objects (Kant 1787, Husserl 1913). The individual experience can have the content that it does only because it resides within that larger unified structure of representation. (See the entry on unity of consciousness.)

4.7 Intentionality and transparency

Conscious mental states are typically regarded as having a representational or intentional aspect in so far as they are about things, refer to things or have satisfaction conditions. One's conscious visual experience correctly represents the world if there are lilacs in a white vase on the table (pace Travis 2004), one's conscious memory is of the attack on the World Trade Center, and one's conscious desire is for a glass of cold water. However, nonconscious states can also exhibit intentionality in such ways, and it is important to understand the ways in which the representational aspects of conscious states resemble and differ from those of nonconscious states (Carruthers 2000). Searle (1990) offers a contrary view according to which only conscious states and dispositions to have conscious states can be genuinely intentional, but most theorists regard intentionality as extending widely into the unconscious domain. (See the entry on consciousness and intentionality.)

One potentially important dimension of difference concerns so called transparency, which is an important feature of consciousness in two interrelated metaphoric senses, each of which has an intentional, an experiential and a functional aspect.

Conscious perceptual experience is often said to be transparent, or in G.E. Moore's (1922) phrase "diaphanous". We transparently "look through" our sensory experience in so far as we seem directly aware of external objects and events present to us rather than being aware of any properties of experience by which it presents or represents such objects to us.

When I look out at the wind-blown meadow, it is the undulating green grass of which I am aware not of any green property of my visual experience. (See the entry on representational theories of consciousness.) Moore himself believed we could become aware of those latter qualities with effort and redirection of attention, though some contemporary transparency advocates deny it (Harman 1990, Tye 1995).

Conscious thoughts and experiences are also transparent in a semantic sense in that their meanings seem immediately known to us in the very act of thinking them (Van Gulick 1992). In that sense we might say to 'think right through' them to what they mean or represent. Transparency in this semantic sense may correspond at least partly with what John Searle calls the "intrinsic intentionality" of consciousness (Searle 1992).

Our conscious mental states seem to have their meanings intrinsically or from the inside just by being what they are in themselves, by contrast with many externalist theories of mental content that ground meaning in causal, counterfactual or informational relations between bearers of intentionality and their semantic or referential objects.

The view of conscious content as intrinsically determined and internally self-evident is sometimes supported by appeals to brain in the vat intuitions, which make it seem that the envatted brain's conscious mental states would keep all their normal intentional contents despite the loss of all their normal causal and informational links to the world (Horgan and Tienson 2002). There is continued controversy about such cases and about competing internalist (Searle 1992) and externalist views (Dretske 1995) of conscious intentionality.

Though semantic transparency and intrinsic intentionality have some affinities, they should not be simply equated, since it may be possible to accommodate the former notion within a more externalist account of content and meaning. Both semantic and sensory transparency obviously concern the representational or intentional aspects of consciousness, but they are also experiential aspects of our conscious life. They are part of what it's like or how it feels phenomenally to be conscious. They also both have functional aspects, in so far as conscious experiences interact with each other in richly content-appropriate ways that manifest our transparent understanding of their contents.

4.8 Dynamic flow

The dynamics of consciousness are evident in the coherent order of its ever changing process of flow and self-transformation, what William James (1890) called the "stream of consciousness." Some temporal sequences of experience are generated by purely internal factors as when one thinks through a puzzle, and others depend in part upon external causes as when one chases a fly ball, but even the latter sequences are shaped in large part by how consciousness transforms itself.

Whether partly in response to outer influences or entirely from within, each moment to moment sequence of experience grows coherently out of those that preceded it,

constrained and enabled by the global structure of links and limits embodied in its underlying prior organization (Husserl 1913). In that respect, consciousness is an autopoietic system, i.e., a self-creating and self-organizing system (Varela and Maturana 1980).

As a conscious mental agent I can do many things such as scan my room, scan a mental image of it, review in memory the courses of a recent restaurant meal along with many of its tastes and scents, reason my way through a complex problem, or plan a grocery shopping trip and execute that plan when I arrive at the market. These are all routine and common activities, but each involves the directed generation of experiences in ways that manifest an implicit practical understanding of their intentional properties and interconnected contents (Van Gulick 2000).

Consciousness is a dynamic process, and thus an adequate descriptive answer to the What question must deal with more than just its static or momentary properties. In particular, it must give some account of the temporal dynamics of consciousness and the ways in which its self-transforming flow reflects both its intentional coherence and the semantic self-understanding embodied in the organized controls through which conscious minds continually remake themselves as autopoietic systems engaged with their worlds.

A comprehensive descriptive account of consciousness would need to deal with more than just these seven features, but having a clear account of each of them would take us a long way toward answering the “What is consciousness?” question.

5. The explanatory question: How can consciousness exist?

The How question focuses on explanation rather than description. It asks us to explain the basic status of consciousness and its place in nature. Is it a fundamental feature of reality in its own right, or does its existence depend upon other nonconscious items, be they physical, biological, neural or computational? And if the latter, can we explain or understand how the relevant nonconscious items could cause or realize consciousness? Put simply, can we explain how to make something conscious out of things that are not conscious?

5.1 Diversity of explanatory projects

The How question is not a single question, but rather a general family of more specific questions (Van Gulick 1995). They all concern the possibility of explaining some sort or aspect of consciousness, but they vary in their particular explananda, the restrictions on their explanans, and their criteria for successful explanation. For example, one might ask whether we can explain access consciousness computationally by mimicking the requisite access relations in a computational model. Or one might be concerned instead with whether the phenomenal and qualitative properties of a conscious creature's mind can be a priori deduced from a description of the neural properties of its brain processes. Both are

versions of the How question, but they ask about the prospects of very different explanatory projects, and thus may differ in their answers (Lycan 1996). It would be impractical, if not impossible, to catalog all the possible versions of the How question, but some of the main options can be listed.

Explananda. Possible explananda would include the various sorts of state and creature consciousness distinguished above, as well as the seven features of consciousness listed in response to the What question. Those two types of explananda overlap and intersect. We might for example aim to explain the dynamic aspect either of phenomenal or of access consciousness. Or we could try to explain the subjectivity of either qualitative or meta-mental consciousness. Not every feature applies to every sort of consciousness, but all apply to several. How one explains a given feature in relation to one sort of consciousness may not correspond with what is needed to explain it relative to another.

Explanans. The range of possible explanans is also diverse. In perhaps its broadest form, the How question asks how consciousness of the relevant sort could be caused or realized by nonconscious items, but we can generate a wealth of more specific questions by further restricting the range of the relevant explanans. One might seek to explain how a given feature of consciousness is caused or realized by underlying neural processes, biological structures, physical mechanisms, functional or teleofunctional relations, computational organization, or even by nonconscious mental states. The prospects for explanatory success will vary accordingly. In general the more limited and elementary the range of the explanans, the more difficult the problem of explaining how could it suffice to produce consciousness (Van Gulick 1995).

Criteria of explanation. The third key parameter is how one defines the criterion for a successful explanation. One might require that the explanandum be a priori deducible from the explanans, although it is controversial whether this is either a necessary or a sufficient criterion for explaining consciousness (Jackson 1993). Its sufficiency will depend in part on the nature of the premises from which the deduction proceeds. As a matter of logic, one will need some bridge principles to connect propositions or sentences about consciousness with those that do not mention it. If one's premises concern physical or neural facts, then one will need some bridge principles or links that connect such facts with facts about consciousness (Kim 1998). Brute links, whether nomic or merely well confirmed correlations, could provide a logically sufficient bridge to infer conclusions about consciousness. But they would probably not allow us to see how or why those connections hold, and thus they would fall short of fully explaining how consciousness exists (Levine 1983, 1993, McGinn 1991).

One could legitimately ask for more, in particular for some account that made intelligible why those links hold and perhaps why they could not fail to do so. A familiar two-stage model for explaining macro-properties in terms of micro-substrates is often invoked. In the first step, one analyzes the macro-property in terms of functional conditions, and then in the second stage one shows that the micro-structures obeying the laws of their own

level nomically suffice to guarantee the satisfaction of the relevant functional conditions (Armstrong 1968, Lewis 1972).

The micro-properties of collections of H₂O molecules at 20°C suffice to satisfy the conditions for the liquidity of the water they compose. Moreover, the model makes intelligible how the liquidity is produced by the micro-properties. A satisfactory explanation of how consciousness is produced might seem to require a similar two stage story. Without it, even a priori deducibility might seem explanatorily less than sufficient, though the need for such a story remains a matter of controversy (Block and Stalnaker 1999, Chalmers and Jackson 2001).

5.2 The explanatory gap

Our current inability to supply a suitably intelligible link is sometimes described, following Joseph Levine (1983), as the existence of an explanatory gap, and as indicating our incomplete understanding of how consciousness might depend upon a nonconscious substrate, especially a physical substrate. The basic gap claim admits of many variations in generality and thus in strength.

In perhaps its weakest form, it asserts a practical limit on our present explanatory abilities; given our current theories and models we can not now articulate an intelligible link. A stronger version makes an in principle claim about our human capacities and thus asserts that given our human cognitive limits we will never be able to bridge the gap. To us, or creatures cognitively like us, it must remain a residual mystery (McGinn 1991). Colin McGinn (1995) has argued that given the inherently spatial nature of both our human perceptual concepts and the scientific concepts we derive from them, we humans are not conceptually suited for understanding the nature of the psychophysical link. Facts about that link are as cognitively closed to us as are facts about multiplication or square roots to armadillos. They do not fall within our conceptual and cognitive repertoire. An even stronger version of the gap claim removes the restriction to our cognitive nature and denies in principle that the gap can be closed by any cognitive agents.

Those who assert gap claims disagree among themselves about what metaphysical conclusions, if any, follow from our supposed epistemic limits. Levine himself has been reluctant to draw any anti-physicalist ontological conclusions (Levine 1993, 2001). On the other hand some neodualists have tried to use the existence of the gap to refute physicalism (Foster 1996, Chalmers 1996). The stronger one's epistemological premise, the better the hope of deriving a metaphysical conclusion. Thus unsurprisingly, dualist conclusions are often supported by appeals to the supposed impossibility in principle of closing the gap.

If one could see on a priori grounds that there is no way in which consciousness could be intelligibly explained as arising from the physical, it would not be a big step to concluding that it in fact does not do so (Chalmers 1996). However, the very strength of such an

epistemological claim makes it difficult to assume with begging the metaphysical result in question. Thus those who wish to use a strong in principle gap claim to refute physicalism must find independent grounds to support it. Some have appealed to conceivability arguments for support, such as the alleged conceivability of zombies molecularly identical with conscious humans but devoid of all phenomenal consciousness (Campbell 1970, Kirk 1974, Chalmers 1996). Other supporting arguments invoke the supposed non-functional nature of consciousness and thus its alleged resistance to the standard scientific method of explaining complex properties (e.g., genetic dominance) in terms of physically realized functional conditions (Block 1980a, Chalmers 1996). Such arguments avoid begging the anti-physicalist question, but they themselves rely upon claims and intuitions that are controversial and not completely independent of one's basic view about physicalism. Discussion on the topic remains active and ongoing.

Our present inability to see any way of closing the gap may exert some pull on our intuitions, but it may simply reflect the limits of our current theorizing rather than an unbridgeable in principle barrier (Dennett 1991). Moreover, some physicalists have argued that explanatory gaps are to be expected and are even entailed by plausible versions of ontological physicalism, ones that treat human agents as physically realized cognitive systems with inherent limits that derive from their evolutionary origin and situated contextual mode of understanding (Van Gulick 1985, 2003; McGinn 1991, Papineau 1995, 2002). On this view, rather than refuting physicalism, the existence of explanatory gaps may confirm it. Discussion and disagreement on these topics remains active and ongoing.

5.3 Reductive and non-reductive explanation

As the need for intelligible linkage has shown, a priori deducibility is not in itself obviously sufficient for successful explanation (Kim 1980), nor is it clearly necessary. Some weaker logical link might suffice in many explanatory contexts. We can sometimes tell enough of a story about how facts of one sort depend upon those of another to satisfy ourselves that the latter do in fact cause or realize the former even if we can not strictly deduce all the former facts from the latter.

Strict intertheoretical deduction was taken as the reductive norm by the logical empiricist account of the unity of science (Putnam and Oppenheim 1958), but in more recent decades a looser nonreductive picture of relations among the various sciences has gained favor. In particular, nonreductive materialists have argued for the so called "autonomy of the special sciences" (Fodor 1974) and for the view that understanding the natural world requires us to use a diversity of conceptual and representational systems that may not be strictly intertranslatable or capable of being put into the tight correspondence required by the older deductive paradigm of interlevel relations (Putnam 1975).

Economics is often cited as an example (Fodor 1974, Searle 1992). Economic facts may be realized by underlying physical processes, but no one seriously demands that we be able to deduce the relevant economic facts from detailed descriptions of their underlying

physical bases or that we be able to put the concepts and vocabulary of economics in tight correspondence with those of the physical sciences.

Nonetheless our deductive inability is not seen as cause for ontological misgivings; there is no “money-matter” problem. All that we require is some general and less than deductive understanding of how economic properties and relations might be underlain by physical ones. Thus one might opt for a similar criterion for interpreting the How question and for what counts as explaining how consciousness might be caused or realized by nonconscious items. However, some critics, such as Kim (1987), have challenged the coherence of any view that aims to be both non-reductive and physicalist, though supporters of such views have replied in turn (Van Gulick 1993).

Others have argued that consciousness is especially resistant to explanation in physical terms because of the inherent differences between our subjective and objective modes of understanding. Thomas Nagel famously argued (1974) that are unavoidable limits placed on our ability to understand the phenomenology of bat experience by our inability to empathetically take on an experiential perspective like that which characterizes the bat's echo-locatory auditory experience of its world. Given our inability to undergo similar experience, we can have at best partial understanding of the nature of such experience. No amount of knowledge gleaned from the external objective third-person perspective of the natural sciences will supposedly suffice to allow us to understand what the bat can understand of its own experience from its internal first-person subjective point of view.

5.4 Prospects of explanatory success

The How question thus subdivides into a diverse family of more specific questions depending upon the specific sort or feature of consciousness one aims to explain, the specific restrictions one places on the range of the explanans and the criterion one uses to define explanatory success. Some of the resulting variants seem easier to answer than others. Progress may seem likely on some of the so called “easy problems” of consciousness, such as explaining the dynamics of access consciousness in terms of the functional or computational organization of the brain (Baars 1988). Others may seem less tractable, especially the so-called “hard problem” (Chalmers 1995) which is more or less that of giving an intelligible account that lets us see in an intuitively satisfying way how phenomenal or “what it's like” consciousness might arise from physical or neural processes in the brain.

Positive answers to some versions of the How questions seem near at hand, but others appear to remain deeply baffling. Nor should we assume that every version has a positive answer. If dualism is true, then consciousness in at least some of its types may be basic and fundamental. If so, we will not be able to explain how it arises from nonconscious items since it simply does not do so.

One's view of the prospects for explaining consciousness will typically depend upon one's

perspective. Optimistic physicalists will likely see current explanatory lapses as merely the reflection of the early stage of inquiry and sure to be remedied in the not too distant future (Dennett 1991, Searle 1992, P. M. Churchland 1995). To dualists, those same impasses will signify the bankruptcy of the physicalist program and the need to recognize consciousness as a fundamental constituent of reality in its own right (Robinson 1982, Foster 1989, 1996, Chalmers 1996). What one sees depends in part on where one stands, and the ongoing project of explaining consciousness will be accompanied by continuing debate about its status and prospects for success.

6. The functional question: Why does consciousness exist?

The functional or Why question asks about the value or role of consciousness and thus indirectly about its origin. Does it have a function, and if so what is it? Does it make a difference to the operation of systems in which it is present, and if so why and how? If consciousness exists as a complex feature of biological systems, then its adaptive value is likely relevant to explaining its evolutionary origin, though of course its present function, if it has one, need not be the same as that it may have had when it first arose. Adaptive functions often change over biological time. Questions about the value of consciousness also have a moral dimension in at least two ways. We are inclined to regard an organism's moral status as at least partly determined by the nature and extent to which it is conscious, and conscious states, especially conscious affective states such as pleasures and pains, play a major role in many of the accounts of value that underlie moral theory (Singer 1975).

As with the What and How questions, the Why question poses a general problem that subdivides into a diversity of more specific inquiries. In so far as the various sorts of consciousness, e.g., access, phenomenal, meta-mental, are distinct and separable — which remains an open question — they likely also differ in their specific roles and values. Thus the Why question may well not have a single or uniform answer.

6.1 Causal status of consciousness

Perhaps the most basic issue posed by any version of the Why question is whether or not consciousness of the relevant sort has any causal impact at all. If it has no effects and makes no causal difference whatsoever, then it would seem unable to play any significant role in the systems or organisms in which it is present, thus undercutting at the outset most inquiries about its possible value. Nor can the threat of epiphenomenal irrelevance be simply dismissed as an obvious non-option, since at least some forms of consciousness have been seriously alleged in the recent literature to lack causal status. (See the entry on epiphenomenalism.) Such worries have been raised especially with regard to qualia and qualitative consciousness (Huxley 1874, Jackson 1982, Chalmers 1996), but challenges have also been leveled against the causal status of other sorts including meta-mental consciousness (Velmans 1991).

Both metaphysical and empirical arguments have been given in support of such claims.

Among the former are those that appeal to intuitions about the conceivability and logical possibility of zombies, i.e., of beings whose behavior, functional organization, and physical structure down to the molecular level are identical to those of normal human agents but who lack any qualia or qualitative consciousness. Some (Kirk 1970, Chalmers 1996) assert such beings are possible in worlds that share all our physical laws, but others deny it (Dennett 1991, Levine 2001). If they are possible in such worlds, then it would seem to follow that even in our world, qualia do not affect the course of physical events including those that constitute our human behaviors. If those events unfold in the same way whether or not qualia are present, then qualia appear to be inert or epiphenomenal at least with respect to events in the physical world. However, such arguments and the zombie intuitions on which they rely are controversial and their soundness remains in dispute (Searle 1992, Yablo 1998, Balog 1999).

Arguments of a far more empirical sort have challenged the causal status of meta-mental consciousness, at least in so far as its presence can be measured by the ability to report on one's mental state. Scientific evidence is claimed to show that consciousness of that sort is neither necessary for any type of mental ability nor does it occur early enough to act as a cause of the acts or processes typically thought to be its effects (Velmans 1991). According to those who make such arguments, the sorts of mental abilities that are typically thought to require consciousness can all be realized unconsciously in the absence of the supposedly required self-awareness.

Moreover, even when conscious self-awareness is present, it allegedly occurs too late to be the cause of the relevant actions rather than their result or at best a joint effect of some shared prior cause (Libet 1985). Self-awareness or meta-mental consciousness according to these arguments turns out to be a psychological after-effect rather than an initiating cause, more like a post facto printout or the result displayed on one's computer screen than like the actual processor operations that produce both the computer's response and its display.

Once again the arguments are controversial, and both the supposed data and their interpretation are subjects of lively disagreement (see Flanagan 1992, and commentaries accompanying Velmans 1991). Though the empirical arguments, like the zombie claims, require one to consider seriously whether some forms of consciousness may be less causally potent than is typically assumed, many theorists regard the empirical data as no real threat to the causal status of consciousness.

If the epiphenomenalists are wrong and consciousness, in its various forms, is indeed causal, what sorts of effects does it have and what differences does it make? How do mental processes that involve the relevant sort of consciousness differ from those that lack it? What function(s) might consciousness play? The following six sections (6.2 — 6.7) discuss some of the more commonly given answers. Though the various functions overlap to some degree, each is distinct, and they differ as well in the sorts of consciousness with which each is most aptly linked.

6.2 Flexible control

Increased flexibility and sophistication of control. Conscious mental processes appear to provide highly flexible and adaptive forms of control. Though unconscious automatic processes can be extremely efficient and rapid, they typically operate in ways that are more fixed and predetermined than those which involve conscious self-awareness (Anderson 1983). Conscious awareness is thus of most importance when one is dealing with novel situations and previously unencountered problems or demands (Penfield 1975, Armstrong 1981).

Standard accounts of skill acquisition stress the importance of conscious awareness during the initial learning phase, which gradually gives way to more automatic processes of the sort that require little attention or conscious oversight (Schneider and Shiffrin 1977). Conscious processing allows for the construction or compilation of specifically tailored routines out of elementary units as well as for the deliberate control of their execution.

There is a familiar tradeoff between flexibility and speed; controlled conscious processes purchase their customized versatility at the price of being slow and effortful in contrast to the fluid rapidity of automatic unconscious mental operations (Anderson 1983). The relevant increases in flexibility would seem most closely connected with the meta-mental or higher-order form of consciousness in so far as the enhanced ability to control processes depends upon greater self-awareness. However, flexibility and sophisticated modes of control may be associated as well with the phenomenal and access forms of consciousness.

6.3 Social coordination

Enhanced capacity for social coordination. Consciousness of the meta-mental sort may well involve not only an increase in self-awareness but also an enhanced understanding of the mental states of other minded creatures, especially those of other members of one's social group (Humphreys 1982). Creatures that are conscious in the relevant meta-mental sense not only have beliefs, motives, perceptions and intentions but understand what it is to have such states and are aware of both themselves and others as having them.

This increase in mutually shared knowledge of each other's minds, enables the relevant organisms to interact, cooperate and communicate in more advanced and adaptive ways. Although meta-mental consciousness is the sort most obviously linked to such a socially coordinative role, narrative consciousness of the kind associated with the stream of consciousness is also clearly relevant in so far as it involves the application to one's own case of the interpretative abilities that derive in part from their social application (Ryle 1949, Dennett 1978, 1992).

6.4 Integrated representation

More unified and densely integrated representation of reality. Conscious experience

presents us with a world of objects independently existing in space and time. Those objects are typically present to us in a multi-modal fashion that involves the integration of information from various sensory channels as well as from background knowledge and memory. Conscious experience presents us not with isolated properties or features but with objects and events situated in an ongoing independent world, and it does so by embodying in its experiential organization and dynamics the dense network of relations and interconnections that collectively constitute the meaningful structure of a world of objects (Kant 1787, Husserl 1913, Campbell 1997).

Of course, not all sensory information need be experienced to have an adaptive effect on behavior. Adaptive non-experiential sensory-motor links can be found both in simple organisms, as well as in some of the more direct and reflexive processes of higher organisms. But when experience is present, it provides a more unified and integrated representation of reality, one that typically allows for more open-ended avenues of response (Lorenz 1977). Consider for example the representation of space in an organism whose sensory input channels are simply linked to movement or to the orientation of a few fixed mechanisms such as those for feeding or grabbing prey, and compare it with that in an organism capable of using its spatial information for flexible navigation of its environment and for whatever other spatially relevant aims or goals it may have, as when a person visually scans her office or her kitchen (Gallistel 1990).

It is representation of this latter sort that is typically made available by the integrated mode of presentation associated with conscious experience. The unity of experienced space is just one example of the sort of integration associated with our conscious awareness of an objective world. (See the entry on unity of consciousness.)

This integrative role or value is most obviously associated with the phenomenal sort of consciousness, but in so far as qualia play a role in the presentation of objects in experience this unifying role is also linked with consciousness in the qualitative and what it's like senses as well. Moreover, it is intimately tied to the transparency of experience described in response to the What question, especially to semantic transparency (Van Gulick 1993).

6.5 Informational access

More global informational access. The information carried in conscious mental states is typically available for use by a diversity of mental subsystems and for application to a wide range of potential situations and actions (Baars 1988). Nonconscious information is more likely to be encapsulated within particular mental modules and available for use only with respect to the applications directly connected to that subsystem's operation (Fodor 1983). Making information conscious typically widens the sphere of its influence and the range of ways in which it can be used to adaptively guide or shape both inner and outer behavior. A state's being conscious may be in part a matter of what Dennett calls "cerebral celebrity", i.e., of its ability to have a content-appropriate impact on other mental states.

This particular role is most directly and definitionally tied to the notion of access consciousness (Block 1995), but meta-mental consciousness as well as the phenomenal and qualitative forms all seem plausibly linked to such increases in the availability of information (Armstrong 1981, Tye 1985).

6.6 Freedom of will

Increased freedom of choice or free will. The issue of free will remains a perennial philosophical problem, not only with regard to whether or not it exists but even as to what it might or should consist in (Dennett 1984, van Inwagen 1983, Hasker 1999, Wegner 2002). (See the entry on free will.) The notion of free will may itself remain too murky and contentious to shed any clear light on the role of consciousness, but there is a traditional intuition that the two are deeply linked.

Consciousness has been thought to open a realm of possibilities, a sphere of options within which the conscious self might choose or act freely. At a minimum, consciousness might seem a necessary precondition for any such freedom or self-determination (Hasker 1999). How could one engage in the requisite sort of free choice, while remaining solely within the unconscious domain? How can one determine one's own will without being conscious of it and of the options one has to shape it.

The freedom to choose one's actions and the ability to determine one's own nature and future development may admit of many interesting variations and degree rather than being a simple all or nothing matter, and various forms or levels of consciousness might be correlated with corresponding degrees or types of freedom and self-determination (Dennett 1984, 2003). The link with freedom seems strongest for the meta-mental form of consciousness given its emphasis on self-awareness, but potential connections also seem possible for most of the other sorts as well.

6.7 Intrinsic motivation

Intrinsically motivating states. At least some conscious states appear to have the motive force they do intrinsically. In particular, the functional and motivational roles of conscious affective states, such as pleasures and pains, seem intrinsic to their experiential character and inseparable from their qualitative and phenomenal properties, though the view has been challenged (Nelkin 1989, Rosenthal 1991). The attractive positive motivational aspect of a pleasure seems a part of its directly experienced phenomenal feel, as does the negative affective character of a pain, at least in the case of normal non-pathological experience.

There is considerable disagreement about the extent to which the feel and motive force of pain can dissociate in abnormal cases, and some have denied the existence of such intrinsically motivating aspects all together (Dennett 1991). However, at least in the normal case the negative motivational force of pain seems built right into the feel of the experience

itself.

Just how this might be so remains less than clear, and perhaps the appearance of intrinsic and directly experienced motivational force is illusory. But if it is real, then it may be one of the most important and evolutionarily oldest respects in which consciousness makes a difference to the mental systems and processes in which it is present (Humphreys 1992).

Other suggestions have been made about the possible roles and value of consciousness, and these six surely do not exhaust the options. Nonetheless, they are among the most prominent recent hypotheses, and they provide a fair survey of the sorts of answers that have been offered to the Why question by those who believe consciousness does indeed make a difference.

6.8 Constitutive and contingent roles

One further point requires clarification about the various respects in which the proposed functions might answer the Why question. In particular one should distinguish between constitutive cases and cases of contingent realization. In the former, fulfilling the role constitutes being conscious in the relevant sense, while in the latter case consciousness of a given sort is just one way among several in which the requisite role might be realized (Van Gulick 1993).

For example, making information globally available for use by a wide variety of subsystems and behavioral applications may constitute its being conscious in the access sense. By contrast, even if the qualitative and phenomenal forms of consciousness involve a highly unified and densely integrated representation of objective reality, it may be possible to produce representations having those functional characteristics but which are not qualitative or phenomenal in nature.

The fact that in us the modes of representation with those characteristics also have qualitative and phenomenal properties may reflect contingent historical facts about the particular design solution that happened to arise in our evolutionary ancestry. If so, there may be quite other means of achieving a comparable result without qualitative or phenomenal consciousness. Whether this is the right way to think about phenomenal and qualitative conscious is unclear; perhaps the tie to unified and densely integrated representation is in fact as intimate and constitutive as it seems to be in the case of access consciousness (Carruthers 2000). Regardless of how that issue gets resolved, it is important to not to conflate constitution accounts with contingent realization accounts when addressing the function of consciousness and answering the question of why it exists (Chalmers 1996).

7. Theories of consciousness.

In response to the What, How and Why questions many theories of consciousness have

been proposed in recent years. However, not all theories of consciousness are theories of the same thing. They vary not only in the specific sorts of consciousness they take as their object, but also in their theoretical aims.

Perhaps the largest division is between general metaphysical theories that aim to locate consciousness in the overall ontological scheme of reality and more specific theories that offer detailed accounts of its nature, features and role. The line between the two sorts of theories blurs a bit, especially in so far as many specific theories carry at least some implicit commitments on the more general metaphysical issues. Nonetheless, it is useful to keep the division in mind when surveying the range of current theoretical offerings.

8. Metaphysical theories of consciousness

General metaphysical theories offer answers to the conscious version of the mind-body problem, “What is the ontological status of consciousness relative to the world of physical reality?” The available responses largely parallel the standard mind-body options including the main versions of dualism and physicalism.

8.1 Dualist theories

Dualist theories regard at least some aspects of consciousness as falling outside the realm of the physical, but specific forms of dualism differ in just which aspects those are. (See the entry on dualism.)

Substance dualism, such as traditional Cartesian dualism (Descartes 1644), asserts the existence of both physical and non-physical substances. Such theories entail the existence of non-physical minds or selves as entities in which consciousness inheres. Though substance dualism is at present largely out of favor, it does have some contemporary proponents (Swinburne 1986, Foster 1989, 1996).

Property dualism in its several versions enjoys a greater level of current support. All such theories assert the existence of conscious properties that are neither identical with nor reducible to physical properties but which may nonetheless be instantiated by the very same things that instantiate physical properties. In that respect they might be classified as dual aspect theories. They take some parts of reality — organisms, brains, neural states or processes — to instantiate properties of two distinct and disjoint sorts: physical ones and conscious, phenomenal or qualitative ones. Dual aspect or property dualist theories can be of at least three different types.

Fundamental property dualism regards conscious mental properties as basic constituents of reality on a par with fundamental physical properties such as electromagnetic charge. They may interact in causal and law-like ways with other fundamental properties such as those of physics, but ontologically their existence is not dependent upon nor derivative from any other properties (Chalmers 1996).

Emergent property dualism treats conscious properties as arising from complex organizations of physical constituents but as doing so in a radical way such that the emergent result is something over and above its physical causes and is not a priori predictable from nor explicable in terms of their strictly physical natures. The coherence of such emergent views has been challenged (Kim 1998) but they have supporters (Hasker 1999).

Neutral monist property dualism treats both conscious mental properties and physical properties as in some way dependent upon and derivative from a more basic level of reality, that in itself is neither mental nor physical (Russell 1927, Strawson 1994). However, if one takes dualism to be a claim about there being two distinct realms of fundamental entities or properties, then perhaps neutral monism should not be classified as a version of property dualism in so far as it does not regard either mental or physical properties as ultimate or fundamental.

Panpsychism might be regarded as a fourth type of property dualism in that it regards all the constituents of reality as having some psychic, or at least proto-psychic, properties distinct from whatever physical properties they may have (Nagel 1979). Indeed neutral monism might be consistently combined with some version of panprotopsychism (Chalmers 1996) according to which the proto-mental aspects of micro-constituents can give rise under suitable conditions of combination to full blown consciousness. (See the entry on panpsychism.)

The nature of the relevant proto-psychic aspect remains unclear, and such theories face a dilemma if offered in hope of answering the Hard Problem. Either the proto-psychic properties involve the sort of qualitative phenomenal feel that generates the Hard Problem or they do not. If they do, it is difficult to understand how they could possibly occur as ubiquitous properties of reality. How could an electron or a quark have any such experiential feel? However, if the proto-psychic properties do not involve any such feel, it is not clear how they are any better able than physical properties to account for qualitative consciousness in solving the Hard Problem.

A variety of arguments have been given in favor of dualist and other anti-physicalist theories of consciousness. Some are largely a priori in nature such as those that appeal to the supposed conceivability of zombies (Kirk 1970, Chalmers 1996) or versions of the knowledge argument (Jackson 1982, 1986) which aim to reach an anti-physicalist conclusion about the ontology of consciousness from the apparent limits on our ability to fully understand the qualitative aspects of conscious experience through third-person physical accounts of the brain processes. (See Jackson 1998, 2004 for a contrary view; see also entries on Zombies, and Qualia: The Knowledge Argument) Other arguments for dualism are made on more empirical grounds, such as those that appeal to supposed causal gaps in the chains of physical causation in the brain (Eccles and Popper 1977) or those based on alleged anomalies in the temporal order of conscious awareness (Libet

1982, 1985). Dualist arguments of both sorts have been much disputed by physicalists (P.S. Churchland 1981, Dennett and Kinsbourne 1992).

8.2 Physicalist theories

Most other metaphysical theories of consciousness are versions of physicalism of one familiar sort or another.

Eliminativist theories reductively deny the existence of consciousness or at least the existence of some of its commonly accepted sorts or features. (See the entry on eliminative materialism.) The radical eliminativists reject the very notion of consciousness as muddled or wrong headed and claim that the conscious/nonconscious distinction fails to cut mental reality at its joints (Wilkes 1984, 1988). They regard the idea of consciousness as sufficiently off target to merit elimination and replacement by other concepts and distinctions more reflective of the true nature of mind (P. S. Churchland 1983).

Most eliminativists are more qualified in their negative assessment. Rather than rejecting the notion outright, they take issue only with some of the prominent features that it is commonly thought to involve, such as qualia (Dennett 1990, Carruthers 2000), the conscious self (Dennett 1992), or the so called "Cartesian Theater" where the temporal sequence of conscious experience gets internally projected (Dennett and Kinsbourne 1992). More modest eliminativists, like Dennett, thus typically combine their qualified denials with a positive theory of those aspects of consciousness they take as real, such as the Multiple Drafts Model (section 9.3 below).

Identity theory, at least strict psycho-physical type-type identity theory, offers another strongly reductive option by identifying conscious mental properties, states and processes with physical ones, most typically of a neural or neurophysiological nature. If having a qualitative conscious experience of phenomenal red just is being in a brain state with the relevant neurophysiological properties, then such experiential properties are real but their reality is a straight forwardly physical reality.

Type-type identity theory is so called because it identifies mental and physical types or properties on a par with identifying the property of being water with the property of being composed of H₂O molecules. After a brief period of popularity in the early days of contemporary physicalism during the 1950s and 60s (Place 1956, Smart 1959) it has been far less widely held because of problems such as the multiple realization objection according to which mental properties are more abstract and thus capable of being realized by many diverse underlying structural or chemical substrates (Fodor 1974, Hellman and Thompson 1975). If one and the same conscious property can be realized by different neurophysiological (or even non-neurophysiological) properties in different organisms, then the two properties can not be strictly identical.

Nonetheless the type-type identity theory has enjoyed a recent if modest resurgence at

least with respect to qualia or qualitative conscious properties. This has been in part because treating the relevant psycho-physical link as an identity is thought by some to offer a way of dissolving the explanatory gap problem (Hill and McLaughlin 1998, Papineau 1995, 2003). They argue that if the conscious qualitative property and the neural property are identical, then there is no need to explain how the latter causes or gives rise to the former. It does not cause it, it is it. And thus there is no gap to bridge, and no further explanation is needed. Identities are not the sort of thing that can be explained, since nothing is identical with anything but itself, and it makes no sense to ask why something is identical with itself.

However, others contend that the appeal to type-type identity does not so obviously void the need for explanation (Levine 2001). Even if two descriptions or concepts in fact refer to one and the same property, one may still reasonably expect some explanation of that convergence, some account of how they pick out one and the same thing despite not initially or intuitively seeming to do so. In other cases of empirically discovered property identities, such as that of heat and kinetic energy, there is a story to be told that explains the co-referential convergence, and it seems fair to expect the same in the psycho-physical case. Thus appealing to type-type identities may not in itself suffice to dissolve the explanatory gap problem.

Most physicalist theories of consciousness are neither eliminativist nor based on strict type-type identities. They acknowledge the reality of consciousness but aim to locate it within the physical world on the basis of some psycho-physical relation short of strict property identity.

Among the common variants are those that take conscious reality to supervene on the physical, be composed of the physical, or be realized by the physical.

Functionalist theories in particular rely heavily on the notion of realization to explicate the relation between consciousness and the physical. According to functionalism, a state or process counts as being of a given mental or conscious type in virtue of the functional role it plays within a suitably organized system (Block 1980a). A given physical state realizes the relevant conscious mental type by playing the appropriate role within the larger physical system that contains it. (See the entry on functionalism.) The functionalist often appeals to analogies with other inter-level relations, as between the biological and biochemical or the chemical and the atomic. In each case properties or facts at one level are realized by complex interactions between items at an underlying level.

Critics of functionalism often deny that consciousness can be adequately explicated in functional terms (Block 1980a, 1980b, Levine 1983, Chalmers 1996). According to such critics, consciousness may have interesting functional characteristics but its nature is not essentially functional. Such claims are sometimes supported by appeal to the supposed possibility of absent or inverted qualia, i.e., the possibility of beings who are functionally equivalent to normal humans but who have reversed qualia or none at all. The status of

such possibilities is controversial (Shoemaker 1981, Dennett 1990, Carruthers 2000), but if accepted they would seem to pose a problem for the functionalist. (See the entry on qualia.)

Those who ground ontological physicalism on the realization relation often combine it with a nonreductive view at the conceptual or representational level that stresses the autonomy of the special sciences and the distinct modes of description and cognitive access they provide.

Non-reductive physicalism of this sort denies that the theoretical and conceptual resources appropriate and adequate for dealing with facts at the level of the underlying substrate or realization level must be adequate as well for dealing with those at the realized level (Putnam 1975, Boyd 1980). As noted above in response to the How question, one can believe that all economic facts are physically realized without thinking that the resources of the physical sciences provide all the cognitive and conceptual tools we need for doing economics (Fodor 1974).

Nonreductive physicalism has been challenged for its alleged failure to “pay its physicalist dues” in reductive coin. It is faulted for supposedly not giving an adequate account of how conscious properties are or could be realized by underlying neural, physical or functional structures or processes (Kim 1987, 1998). Indeed it has been charged with incoherence because of its attempt to combine a claim of physical realization with the denial of the ability to spell out that relation in a strict and a priori intelligible way (Jackson 2004).

However, as noted above in discussion of the How question, nonreductive physicalists reply by agreeing that some account of psycho-physical realization is indeed needed, but adding that the relevant account may fall far short of a priori deducibility, yet still suffice to satisfy our legitimate explanatory demands (McGinn 1991, Van Gulick 1985). The issue remains under debate.

9. Specific Theories of Consciousness

Although there are many general metaphysical/ontological theories of consciousness, the list of specific detailed theories about its nature is even longer and more diverse. No brief survey could be close to comprehensive, but six main types of theories may help to indicate the basic range of options: higher-order theories, representational theories, cognitive theories, neural theories, quantum theories and nonphysical theories.

9.1 Higher-order theories

Higher-order (HO) theories analyze the notion of a conscious mental state in terms of reflexive meta-mental self-awareness. The core idea is that what makes a mental state *M* a conscious mental state is the fact that it is accompanied by a simultaneous and non-

inferential higher-order (i.e., meta-mental) state whose content is that one is now in M. Having a conscious desire for some chocolate involves being in two mental states; one must have both a desire for some chocolate and also a higher-order state whose content is that one is now having just such a desire. Unconscious mental states are unconscious precisely in that we lack the relevant higher-order states about them. Their being unconscious consists in the fact that we are not reflexively and directly aware of being in them. (See the entry on higher-order theories of consciousness.)

Higher-order theories come in two main variants that differ concerning the psychological mode of the relevant conscious-making meta-mental states. Higher-order thought (HOT) theories take the required higher-order state to be an assertoric thought-like meta-state (Rosenthal 1986, 1993). Higher-order perception (HOP) theories take them to be more perception-like and associated with a kind of inner sense and intra-mental monitoring systems of some sort (Armstrong 1981, Lycan 1987, 1996).

Each has its relative strengths and problems. HOT theorists note that we have no organs of inner sense and claim that we experience no sensory qualities other than those presented to us by outer directed perception. HOP theorists on the other hand can argue that their view explains some of the additional conditions required by HO accounts as natural consequences of the perception-like nature of the relevant higher-order states. In particular the demands that the conscious-making meta-state be noninferential and simultaneous with its lower level mental object might be explained by the parallel conditions that typically apply to perception. We perceive what is happening now, and we do so in a way that involves no inferences, at least not any explicit personal-level inferences. Those conditions are no less necessary on the HOT view but are left unexplained by it, which might seem to give some explanatory advantage to the HOP model (Lycan 2004, Van Gulick 2000), though some HOT theorists argue otherwise (Carruthers 2000).

Whatever their respective merits, both HOP and HOT theories face some common challenges, including what might be called the generality problem. Having a thought or perception of a given item X — be it a rock, a pen or a potato — does not in general make X a conscious X. Seeing or thinking of the potato on the counter does not make it a conscious potato. Why then should having a thought or perception of a given desire or a memory make it a conscious desire or memory (Dretske 1995, Byrne 1997). Nor will it suffice to note that we do not apply the term “conscious” to rocks or pens that we perceive or think of, but only to mental states that we perceive or think of (Lycan 1997, Rosenthal 1997). That may be true, but what is needed is some account of why it is appropriate to do so.

The higher-order view is most obviously relevant to the meta-mental forms of consciousness, but some of its supporters take it to explain other types of consciousness as well, including the more subjective what it's like and qualitative types. One common strategy is to analyze qualia as mental features that are capable of occurring unconsciously; for example they might be explained as properties of inner states whose

structured similarity relations given rise to beliefs about objective similarities in the world (Shoemaker 1975, 1990). Though unconscious qualia can play that functional role, there need be nothing that it is like to be in a state that has them (Nelkin 1989, Rosenthal 1991, 1997). According to the HO theorist, what-it's-likeness enters only when we become aware of that first-order state and its qualitative properties by having an appropriate meta-state directed at it.

Critics of the HO view have disputed that account, and some have argued that the notion of unconscious qualia on which it relies is incoherent (Papineau 2002). Whether or not such proposed HO accounts of qualia are successful, it is important to note that most HO advocates take themselves to be offering a comprehensive theory of consciousness, or at least the core of such a general theory, rather than merely one limited to some special meta-mental forms of it.

Other variants of HO theory go beyond the standard HOT and HOP versions including some that analyze consciousness in terms of dispositional rather than occurrent higher-order thoughts (Carruthers 2000). Others appeal to implicit rather than explicit higher-order understanding and weaken or remove the standard assumption that the meta-state must be distinct and separate from its lower-order object (Gennaro 1995, Van Gulick 2000, 2004). Other variants of HO theory continue to be offered, and debate between supporters and critics of the basic approach remains active. (See the recent papers in Gennaro 2004.)

9.2 Representationalist theories

Almost all theories of consciousness regard it as having representational features, but so called representationalist theories are defined by the stronger view that its representational features exhaust its mental features (Harman 1990, Tye 1995, 2000). According to the representationalist, conscious mental states have no mental properties other than their representational properties. Thus two conscious or experiential states that share all their representational properties will not differ in any mental respect.

The exact force of the claim depends on how one interprets the idea of being “representationally the same” for which there are many plausible alternative criteria. One could define it coarsely in terms of satisfaction or truth conditions, but understood in that way the representationalist thesis seems clearly false. There are too many ways in which states might share their satisfaction or truth conditions yet differ mentally, including those that concern their mode of conceptualizing or presenting those conditions.

At the opposite extreme, one could count two states as representationally distinct if they differed in any features that played a role in their representational function or operation. On such a liberal reading any differences in the bearers of content would count as representational differences even if they bore the same intentional or representational content; they might differ only in their means or mode of representation not their content.

Such a reading would of course increase the plausibility of the claim that a conscious state's representational properties exhaust its mental properties but at the cost of significantly weakening or even trivializing the thesis. Thus the representationalist seems to need an interpretation of representational sameness that goes beyond mere satisfaction conditions and reflects all the intentional or contentful aspects of representation without being sensitive to mere differences in underlying non-contentful features of the processes at the realization level. Thus most representationalists provide conditions for conscious experience that include both a content condition plus some further causal role or format requirements (Tye 1995, Dretske 1995, Carruthers 2000). Other representationalists accept the existence of qualia but treat them as objective properties that external objects are represented as having, i.e., they treat them as represented properties rather than as properties of representations or mental states (Dretske 1995, Lycan 1996).

Representationalism can be understood as a qualified form of eliminativism insofar as it denies the existence of properties of a sort that conscious mental states are commonly thought to have — or at least seem to have — namely those that are mental but not representational. Qualia, at least if understood as intrinsic monadic properties of conscious states accessible to introspection, would seem to be the most obvious targets for such elimination. Indeed part of the motivation for representationalism is to show that one can accommodate all the facts about consciousness, perhaps within a physicalist framework, without needing to find room for qualia or any other apparently non-representational mental properties (Dennett 1990, Lycan 1996, Carruthers 2000).

Representationalism has been quite popular in recent years and had many defenders, but it remains highly controversial and intuitions clash about key cases and thought experiments (Block 1996). In particular the possibility of inverted qualia provides a crucial test case. To anti-representationalists, the mere logical possibility of inverted qualia shows that conscious states can differ in a significant mental respect while coinciding representationally. Representationalists in reply deny either the possibility of such inversion or its alleged import (Dretske 1995, Tye 2000).

Many other arguments have been made for and against representationalism, such as those concerning perceptions in different sense modalities of one and the same state of affairs — seeing and feeling the same cube — which might seem to involve mental differences distinct from how the relevant states represent the world to be (Peacocke 1983, Tye 2003). In each case, both sides can muster strong intuitions and argumentative ingenuity. Lively debate continues.

9.3 Cognitive theories

Models that aim to explain consciousness in terms of cognitive process have been offered both by philosophers and psychologists.

The most prominent philosophical example is the Multiple Drafts Model (MDM) of

consciousness, advanced by Daniel Dennett (1991). It combines elements of both representationalism and higher-order theory but does so in a way that varies interestingly from the more standard versions of either. The MDM includes many distinct but interrelated features.

Its name reflects the fact that at any given moment content fixations of many sorts are occurring throughout the brain. What makes some of these contents conscious is not that they occur in a privileged spatial or functional location — the so called “Cartesian Theater” — nor in a special mode or format, all of which the MDM denies. Rather it is a matter of what Dennett calls “cerebral celebrity”, i.e., the degree to which a given content influences the future development of other contents throughout the brain, especially with regard to how those effects are manifest in the reports and behaviors that the person makes in response to various probes that might indicate her conscious state.

One of the MDM's key claims is that different probes (e. g., being asked different questions or being in different contexts that make differing behavioral demands) may elicit different answers about the person's conscious state. Moreover, according to the MDM there may be no probe-independent fact of the matter about what the person's conscious state really was. Hence the “multiple” of the Multiple Drafts Model.

The MDM is representationalist in that it analyzes consciousness in terms of content relations. It also denies the existence of qualia and thus rejects any attempt to distinguish conscious states from nonconscious states by their presence. It rejects as well the notion of the self as an inner observer, whether located in the Cartesian Theater or elsewhere. The MDM treats the self as an emergent or virtual aspect of the coherent roughly serially narrative that is constructed through the interactive play of contents in the system. Many of those contents are bound together at the intentional level as perceptions or fixations from a relatively unified and temporally extended point of view, i.e., they cohere in their contents as if they were the experiences of a ongoing self. But it is the order of dependence that is crucial to the MDM account. The relevant contents are not unified because they are all observed by a single self, but just the converse. It is because they are unified and coherent at the level of content that they count as the experiences of a single self, at least of a single virtual self.

It is in this respect that the MDM shares some elements with higher-order theories. The contents that compose the serial narrative are at least implicitly those of a ongoing if virtual self, and it is they that are most likely to be expressed in the reports the person makes of her conscious state in response to various probes. They thus involve a certain degree of reflexivity or self-awareness of the sort that is central to higher-order theories, but the higher-order aspect is more an implicit feature of the stream of contents rather than present in distinct explicit higher-order states of the sort found in standard HO theories.

Dennett's MDM has been highly influential but has also drawn criticism, especially from

those who find it insufficiently realist in its view of consciousness and at best incomplete in achieving its stated goal to fully explain it (Block 1994, Dretske 1994, Levine 1994). Many of its critics acknowledge the insight and value of the MDM, but deny that there are no real facts of consciousness other than those captured by it (Rosenthal 1994, Van Gulick 1994, Akins 1996).

A major psychological example of the cognitive approach is the global workspace theory of Bernard Baars (1988). It offers a psychological and largely functional model of consciousness which deals most directly with the access notion of consciousness and has much in common with the multiple draft model's idea of consciousness as a form of cerebral celebrity. The main idea of global workspace theories is that consciousness is a limited resource capacity or module that enables information to be "broadcast" widely throughout the system and allows for more flexible sophisticated processing. It is thus closely allied with many models in cognitive psychology concerned with attention and working memory. Many other psychologically based cognitive models of consciousness have been advanced, including those of George Mandler (1975), Timothy Shallice, (1988) and Daniel Schacter (1989).

9.4 Neural theories

Neural theories of consciousness come in many forms, though most in some way concern the so called "neural correlates of consciousness" or NCCs. Unless one is a dualist or other non-physicalist, more than mere correlation is required; at least some NCCs must be the essential substrates of consciousness. An explanatory neural theory needs to explain why or how the relevant correlations exist, and if the theory is committed to physicalism that will require showing how the underlying neural substrates could be identical with their neural substrates or at least realize them by satisfying the required roles or conditions (Metzinger 2000).

Such theories are diverse not only in the neural processes or properties to which they appeal but also in the aspects of consciousness they take as their respective explananda. Some are based on high-level systemic features of the brain, but others focus on more specific physiological or structural properties, with corresponding differences in their intended explanatory targets.

A sampling of recent neural theories might include models that appeal to global integrated fields (Kinsbourne), binding through synchronous oscillation (Singer 1999, Crick and Koch 1990), NMDA-mediated transient neural assemblies (Flohr 1995), thalamically modulated patterns of cortical activation (Llinas 2001), reentrant cortical loops (Edelman 1989), comparator mechanisms that engage in continuous action-prediction-assessment loops between frontal and midbrain areas (Gray 1995), left hemisphere based interpretative processes (Gazzaniga 1988), and emotive somatosensory hemostatic processes based in the frontal-limbic nexus (Damasio 1999) or in the periaqueductal gray (Panksepp 1998).

In each case the aim is to explain how organization and activity at the relevant neural level could underlie one or another major type or feature of consciousness. Global fields or transient synchronous assemblies could underlie the intentional unity of phenomenal consciousness. NMDA-based plasticity, specific thalamic projections into the cortex, or regular oscillatory waves could all contribute to the formation of short term but widespread neural patterns or regularities needed to knit integrated conscious experience out of the local activity in diverse specialized brain modules. Left hemisphere interpretative processes could provide a basis for narrative forms of conscious self-awareness. Thus it is possible for multiple distinct neural theories to all be true, with each contributing some partial understanding of the links between conscious mentality in its diverse forms and the active brain at its many levels of complex organization and structure.

9.5 Quantum theories

Other physical theories have gone beyond the neural and placed the natural locus of consciousness at a far more fundamental level, in particular at the micro-physical level of quantum phenomena. According to such theories, the nature and basis of consciousness can not be adequately understood within the framework of classical physics but must be sought within the alternative picture of physical reality provided by quantum mechanics. The proponents of the quantum consciousness approach regard the radically alternative and often counterintuitive nature of quantum physics as just what is needed to overcome the supposed explanatory obstacles that confront more standard attempts to bridge the psycho-physical gap.

Again there are a wide range of specific theories and models that have been proposed, appealing to a variety of quantum phenomena to explain a diversity of features of consciousness. It would be impossible to catalog them here or even explain in any substantial way the key features of quantum mechanics to which they appeal. However, a brief selective survey may provide a sense, however partial and obscure, of the options that have been proposed.

The physicist Roger Penrose (1989, 1994) and the anesthesiologist Stuart Hameroff (1998) have championed a model according to which consciousness arises through quantum effects occurring within subcellular structures internal to neurons known as microtubules. The model posits so called “objective collapses” which involve the quantum system moving from a superposition of multiple possible states to a single definite state, but without the intervention of an observer or measurement as in most quantum mechanical models. According to the Penrose and Hameroff, the environment internal to the microtubules is especially suitable for such objective collapses, and the resulting self-collapses produce a coherent flow regulating neuronal activity and making non-algorithmic mental processes possible.

The psychiatrist Ian Marshall has offered a model that aims to explain the coherent unity

of consciousness by appeal to the production within the brain of a physical state akin to that of a Bose-Einstein condensate. The latter is a quantum phenomenon in which a collection of atoms acts as a single coherent entity and the distinction between discrete atoms is lost. While brain states are not literally examples of Bose-Einstein condensates, reasons have been offered to show why brains are likely to give rise to states that are capable of exhibiting a similar coherence (Marshall and Zohar 1990).

A basis for consciousness has also been sought in the holistic nature of quantum mechanics and the phenomenon of entanglement, according to which particles that have interacted continue to have their natures depend upon each other even after their separation. Unsurprisingly these models have been targeted especially at explaining the coherence of consciousness, but they have also been invoked as a more general challenge to the atomistic conception of traditional physics according to which the properties of wholes are to be explained by appeal to the properties of their parts plus their mode of combination, a method of explanation that might be regarded as unsuccessful to date in explaining consciousness (Silberstein 1998, 2001).

Others have taken quantum mechanics to indicate that consciousness is an absolutely fundamental property of physical reality, one that needs to be brought in at the very most basic level (Stapp 1993). They have appealed especially to the role of the observer in the collapse of the wave function, i.e., the collapse of quantum reality from a superposition of possible states to a single definite state when a measurement is made. Such models may or may not embrace a form of quasi-idealism, in which the very existence of physical reality depends upon its being consciously observed.

There are many other quantum models of consciousness to be found in the literature — some advocating a radically revisionist metaphysics and others not — but these four provide a reasonable, though partial, sample of the alternatives.

9.6 Non-physical theories

Most specific theories of consciousness — whether cognitive, neural or quantum mechanical — aim to explain or model consciousness as a natural feature of the physical world. However, those who reject a physicalist ontology of consciousness must find ways of modeling it as a nonphysical aspect of reality. Thus those who adopt a dualist or anti-physicalist metaphysical view must in the end provide specific models of consciousness different from the five types above. Both substance dualists and property dualists must develop the details of their theories in ways that articulate the specific natures of the relevant non-physical features of reality with which they equate consciousness or to which they appeal in order to explain it.

A variety of such models have been proposed including the following. David Chalmers (1996) has offered an admittedly speculative version of panpsychism which appeals to the notion of information not only to explain psycho-physical invariances between

phenomenal and physically realized information spaces but also to possibly explain the ontology of the physical as itself derived from the informational (a version of “it from bit” theory). In a somewhat similar vein, Gregg Rosenberg has recently (2004) proposed an account of consciousness that simultaneously addresses the ultimate categorical basis of causal relations. In both the causal case and the conscious case, Rosenberg argues the relational-functional facts must ultimately depend upon a categorical non-relational base, and he offers a model according to which causal relations and qualitative phenomenal facts both depend upon the same base. Also, as noted just above (section 9.5), some quantum theories treat consciousness as a fundamental feature of reality (Stapp 1993), and insofar as they do so, they might be plausibly classified as non-physical theories as well.

10. Conclusion

A comprehensive understanding of consciousness will likely require theories of many types. One might usefully and without contradiction accept a diversity of models that each in their own way aim respectively to explain the physical, neural, cognitive, functional, representational and higher-order aspects of consciousness. There is unlikely to be any single theoretical perspective that suffices for explaining all the features of consciousness that we wish to understand. Thus a synthetic and pluralistic approach may provide the best road to future progress.

Bibliography

- Akins, K. 1993. “A bat without qualities?” In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Akins, K. 1996. “Lost the plot? Reconstructing Dennett's multiple drafts theory of consciousness.” *Mind and Language*, 11: 1-43.
- Anderson, J. 1983. *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Armstrong, D. 1968. *A Materialist Theory of Mind*, London: Routledge and Kegan Paul.
- Armstrong, D. 1981. “What is consciousness?” In *The Nature of Mind*. Ithaca, NY: Cornell University Press.
- Baars, B. 1988. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Balog, K. 1999. “Conceivability, possibility, and the mind-body problem.” *Philosophical Review*, 108: 497-528.
- Block, N. 1980a. “Troubles with Functionalism,” in *Readings in the Philosophy of Psychology*, Volume 1, Ned Block, ed., Cambridge, MA : Harvard University Press, 268-305.
- Block, N. 1980b. “Are absent qualia impossible?” *Philosophical Review*, 89/2: 257-74.
- Block, N. 1990. “Inverted Earth,” *Philosophical Perspectives*, 4, J. Tomberlin, ed., Atascadero, CA: Ridgeview Publishing Company.
- Block, N. 1995. “On a confusion about the function of consciousness.” *Behavioral and Brain Sciences*, 18: 227-47.
- Block, N. 1994. “What is Dennett's theory a theory of?” *Philosophical Topics*, 22/1-2: 23-

- Block, N. 1996. "Mental paint and mental latex." In E. Villanueva, ed. *Perception*. Atascadero, CA: Ridgeview.
- Block, N. and Stalnaker, R. 1999. "Conceptual analysis, dualism, and the explanatory gap." *Philosophical Review*, 108/1: 1-46.
- Boyd, R. 1980. "Materialism without reductionism: What physicalism does not entail." In N. Block, ed. *Readings in the Philosophy of Psychology*, Vol. 1. Cambridge, MA: Harvard University Press.
- Byrne, A. 1997. "Some like it HOT: consciousness and higher-order thoughts." *Philosophical Studies*, 2: 103-29.
- Byrne, A. 2001. "Intentionalism defended". *Philosophical Review*, 110: 199-240.
- Campbell, K. 1970. *Body and Mind*. New York: Doubleday.
- Campbell, J. 1994. *Past, Space, and Self*. Cambridge, MA: MIT Press.
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge University Press.
- Chalmers, D. 1995. "Facing up to the problem of consciousness". *Journal of Consciousness Studies*, 2: 200-19.
- Chalmers, D. 1996. *The Conscious Mind*. Oxford: Oxford University Press.
- Chalmers, D. 2002. "Does conceivability entail possibility?" In T. Gendler and J. Hawthorne eds. *Conceivability and Possibility*. Oxford: Oxford University Press.
- Chalmers, D. 2003. "The content and epistemology of phenomenal belief." In A. Jokic and Q. Smith eds. *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.
- Chalmers, D. and Jackson, F. 2001. "Conceptual analysis and reductive explanation". *Philosophical Review*, 110/3: 315-60.
- Churchland, P. M. 1985. "Reduction, qualia, and direct introspection of brain states". *Journal of Philosophy*, 82: 8-28.
- Churchland, P. M. 1995. *The Engine of Reason and Seat of the Soul*. Cambridge, MA: MIT Press.
- Churchland, P. S. 1981. "On the alleged backwards referral of experiences and its relevance to the mind body problem". *Philosophy of Science*, 48: 165-81.
- Churchland, P. S. 1983. "Consciousness: the transmutation of a concept". *Pacific Philosophical Quarterly*, 64: 80-95.
- Churchland, P. S. 1996. "The hornswoggle problem". *Journal of Consciousness Studies*, 3: 402-8.
- Clark, A. 1993. *Sensory Qualities*. Oxford: Oxford University Press.
- Clark, G. and Riel-Salvatore, J. 2001. "Grave markers, middle and early upper paleolithic burials". *Current Anthropology*, 42/4: 481-90.
- Cleeremans, A., ed. 2003. *The Unity of Consciousness: Binding, Integration and Dissociation*. Oxford: Oxford University Press.
- Crick, F. and Koch, C. 1990. "Toward a neurobiological theory of consciousness". *Seminars in Neuroscience*, 2: 263-75.
- Crick, F. H. 1994. *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Scribners.

- Davies, M. and Humphreys, G. 1993. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Damasio, A. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt.
- Dennett, D. C. 1978. *Brainstorms*. Cambridge: MIT Press.
- Dennett, D. C. 1984. *Elbow Room: The Varieties of Free Will Worth Having*. Cambridge: MIT Press.
- Dennett, D. C. 1990. "Quining qualia". In *Mind and Cognition*, W. Lycan, ed., Oxford : Blackwell, 519-548.
- Dennett, D. C. 1991. *Consciousness Explained*. Boston: Little, Brown and Company.
- Dennett, D. C. 1992. "The self as the center of narrative gravity". In F. Kessel, P. Cole, and D. L. Johnson, eds. *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ: Lawrence Erlbaum.
- Dennett, D. C. 2003. *Freedom Evolves*. New York: Viking.
- Dennett, D. C. and Kinsbourne, M. 1992. "Time and the observer: the where and when of consciousness in the brain". *Behavioral and Brain Sciences*, 15: 187-247.
- Descartes, R. 1644/1911. *The Principles of Philosophy*. Translated by E. Haldane and G. Ross. Cambridge: Cambridge University Press.
- Dretske, F. 1993. "Conscious experience." *Mind*, 102: 263-283.
- Dretske, F. 1994. "Differences that make no difference". *Philosophical Topics*, 22/1-2: 41-58.
- Dretske, F. 1995. *Naturalizing the Mind*. Cambridge, Mass: The MIT Press, Bradford Books.
- Eccles, J. and Popper, K. 1977. *The Self and Its Brain: An Argument for Interactionism*. Berlin: Springer
- Edelman, G. 1989. *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books.
- Farah, M. 1990. *Visual Agnosia*. Cambridge: MIT Press.
- Flanagan, O. 1992. *Consciousness Reconsidered*. Cambridge, MA: MIT Press.
- Flohr, H. 1995. "An information processing theory of anesthesia". *Neuropsychologia*, 33/9: 1169-80.
- Flohr, H., Glade, U. and Motzko, D. 1998. "The role of the NMDA synapse in general anesthesia". *Toxicology Letters*, 100-101: 23-29.
- Fodor, J. 1974. "Special sciences". *Synthese*, 28: 77-115.
- Fodor, J. 1983. *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Foster, J. 1989. "A defense of dualism". In J. Smythies and J. Beloff, eds. *The Case for Dualism*. Charlottesville, VA: University of Virginia Press.
- Foster J. 1996. *The Immaterial Self: A Defence of the Cartesian Dualist Conception of Mind*. London: Routledge.
- Gallistel, C. 1990. *The Organization of Learning*. Cambridge, MA: MIT Press.
- Gardiner, H. 1985. *The Mind's New Science*. New York: Basic Books.
- Gazzaniga, M. 1988. *Mind Matters: How Mind and Brain Interact to Create our Conscious Lives*. Boston: Houghton Mifflin.
- Gennaro, R. 1995. *Consciousness and Self-consciousness: A Defense of the Higher-Order*

Thought Theory of Consciousness. Amsterdam and Philadelphia: John Benjamins.

- Gennaro, R., ed. 2004. Higher-Order Theories of Consciousness. Amsterdam and Philadelphia: John Benjamins.

- Gray, J. 1995. "The contents of consciousness: a neuropsychological conjecture". Behavior and Brain Sciences, 18/4: 659-722.

- Hameroff, S. 1998. "Quantum computation in brain microtubules? The Penrose-Hameroff "Orch OR" model of consciousness". Philosophical Transactions Royal Society London, A 356: 1869-96.

- Hardin, C. 1986. Color for Philosophers. Indianapolis: Hackett.

- Hardin, C. 1992. "Physiology, phenomenology, and Spinoza's true colors". In A. Beckermann, H. Flohr, and J. Kim, eds. Emergence or Reduction?: Prospects for Nonreductive Physicalism. Berlin and New York: De Gruyter.

- Harman, G. 1990. "The intrinsic quality of experience". In J. Tomberlin, ed. Philosophical Perspectives, 4. Atascadero, CA: Ridgeview Publishing .

- Hartshorne, C. 1978. "Panpsychism: mind as sole reality". Ultimate Reality and Meaning, 1: 115-29.

- Hasker, W. 1999. The Emergent Self. Ithaca, NY: Cornell University Press.

- Heidegger, M. 1927/1962. Being and Time (Sein und Zeit). Translated by J. Macquarrie and E. Robinson. New York: Harper and Row.

- Hellman, G. and Thompson, F. 1975. "Physicalism: ontology, determination and reduction". Journal of Philosophy, 72: 551-64.

- Hill, C. 1991. Sensations: A Defense of Type Materialism. Cambridge: Cambridge University Press.

- Hill, C. 1997. "Imaginability, conceivability, possibility, and the mind-body problem". Philosophical Studies, 87: 61-85.

- Hill, C. and McLaughlin, B. 1998. "There are fewer things in reality than are dreamt of in Chalmers' philosophy". Philosophy and Phenomenological Research, 59/2: 445-54.

- Horgan, T. 1984. "Jackson on Physical Information and Qualia." Philosophical Quarterly, 34: 147-83.

- Horgan, T. and Tienson, J. 2002. "The intentionality of phenomenology and the phenomenology of intentionality". In D. J. Chalmers , ed., Philosophy of Mind: Classical and Contemporary Readings. New York: Oxford University Press.

- Hume, D. 1739/1888. A Treatise of Human Nature. ed. L Selby-Bigge. Oxford: Oxford University Press.

- Humphreys, N. 1982. Consciousness Regained. Oxford: Oxford University Press.

- Humphreys, N. 1992. A History of the Mind. London: Chatto and Windus.

- Husserl, E. 1913/1931. Ideas: General Introduction to Pure Phenomenology (Ideen au einer reinen Phänomenologie und phänomenologischen Philosophie). Translated by W. Boyce Gibson. New York: MacMillan.

- Husserl, E. 1929/1960. Cartesian Meditations: an Introduction to Phenomenology. Translated by Dorian Cairns. The Hague: M. Nijhoff.

- Huxley, T. 1866. Lessons on Elementary Physiology 8. London

- Huxley, T. 1874. "On the hypothesis that animals are automata". Fortnightly Review, 95: 555-80. Reprinted in Collected Essays. London, 1893.

- Hurley, S. 1998. *Consciousness in Action*. Cambridge, MA: Harvard University Press.
- Jackson, F. 1982. "Epiphenomenal qualia". *Philosophical Quarterly*, 32: 127-136.
- Jackson, F. 1986. "What Mary didn't know". *Journal of Philosophy*, 83: 291-5.
- Jackson, F. 1993. "Armchair metaphysics". In J. O'Leary-Hawthorne and M. Michael, eds. *Philosophy of Mind*. Dordrecht: Kluwer Books.
- Jackson, F. 1998. "Postscript on qualia". In F. Jackson *Mind, Method and Conditionals*. London: Routledge.
- Jackson, F. 2004. "Mind and illusion." In P. Ludlow, Y. Nagasawa and D. Stoljar eds. *There's Something about Mary: Essays on the Knowledge Argument*. Cambridge, MA: MIT Press.
- James, W. 1890. *The Principles of Psychology*. New York: Henry Holt and Company.
- Jaynes, J. 1974. *The Origins of Consciousness in the Breakdown of the Bicameral Mind*. Boston: Houghton Mifflin.
- Kim, J. 1987. "The myth of non-reductive physicalism". *Proceedings and Addresses of the American Philosophical Association*.
- Kim, J. 1998. *Mind in Physical World*. Cambridge: MIT Press.
- Kant, I. 1787/1929. *Critique of Pure Reason*. Translated by N. Kemp Smith. New York: MacMillan.
- Kinsbourne, M. 1988. "Integrated field theory of consciousness". In A. Marcel and E. Bisiach, eds. *Consciousness in Contemporary Science*. Oxford: Oxford University Press.
- Kirk, R. 1974. "Zombies vs materialists". *Proceedings of the Aristotelian Society*, Supplementary Volume, 48: 135-52.
- Kirk, R. 1991. "Why shouldn't we be able to solve the mind-body problem?" *Analysis*, 51: 17-23.
- Köhler, W. 1929. *Gestalt Psychology*. New York: Liveright.
- Köffka, K. 1935. *Principles of Gestalt Psychology*. New York: Harcourt Brace.
- Leibniz, G. W. 1686 /1991. *Discourse on Metaphysics*. Translated by D. Garter and R. Aries. Indianapolis: Hackett.
- Leibniz, G. W. 1720/1925. *The Monadology*. Translated by R. Lotte. London: Oxford University Press.
- Levine, J. 1983. "Materialism and qualia: the explanatory gap". *Pacific Philosophical Quarterly*, 64: 354-361.
- Levine, J. 1993. "On leaving out what it's like". In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Levine, J. 1994. "Out of the closet: a qualophile confronts qualophobia". *Philosophical Topics*, 22/1-2: 107-26.
- Levine, J. 2001. *Purple Haze: The Puzzle of Conscious Experience*. Cambridge, Mass: The MIT Press.
- Lewis, D. 1972. "Psychophysical and theoretical identifications". *Australasian Journal of Philosophy*, 50: 249-58.
- Lewis, D. 1990. "What experience teaches." In W. Lycan, ed. *Mind and Cognition: A Reader*. Oxford: Blackwell.
- Libet, B. 1982. "Subjective antedating of a sensory experience and mind-brain theories". *Journal of Theoretical Biology*, 114: 563-70.

- Libet, B. 1985. "Unconscious cerebral initiative and the role of conscious will in voluntary action". Behavioral and Brain Sciences, 8: 529-66.
- Llinas, R. 2001. I of the vortex: from neurons to self. Cambridge, MA: MIT Press
- Loar, B. 1990. "Phenomenal states," in Philosophical Perspectives, 4: 81-108.
- Loar, B. 1997. "Phenomenal states". In N. Block, O. Flanagan, and G. Guzeldere eds. The Nature of Consciousness. Cambridge, MA: MIT Press.
- Locke, J. 1688/1959. An Essay on Human Understanding. New York: Dover.
- Lockwood, M. 1989. Mind, Brain, and the Quantum. Oxford: Oxford University Press.
- Lorenz, K. 1977. Behind the Mirror (Rückseite dyes Speigels). Translated by R. Taylor. New York: Harcourt Brace Jovanovich.
- Lycan, W. 1987. Consciousness. Cambridge, MA: MIT Press.
- Lycan, W. 1996. Consciousness and Experience. Cambridge, MA: MIT Press.
- Lycan, W. 2004. "The superiority of HOP to HOT". In R. Gennaro ed. Higher-Order Theories of Consciousness. Amsterdam and Philadelphia: John Benjamins.
- Mandler, G. 1975. Mind and Emotion. New York: Wiley.
- Marshall, I. and Zohar, D. 1990. The Quantum Self: Human Nature and Consciousness Defined by the New Physics. New York: Morrow.
- McGinn, C. 1989. "Can we solve the mind-body problem?" Mind, 98: 349-66
- McGinn, C. 1991. The Problem of Consciousness. Oxford: Blackwell.
- McGinn, C. 1995. "Consciousness and space." In T. Metzinger, ed. Conscious Experience. Paderborn: Ferdinand Schöningh.
- Merleau-Ponty, M. 1945/1962. Phenomenology of Perception (Phénoménologie de l'ye Perception). Translated by Colin Smith. London: Routledge and Kegan Paul.
- Metzinger, T., ed. 1995. Conscious Experience. Paderborn: Ferdinand Schöningh.
- Metzinger, T. ed. 2000. Neural Correlates of Consciousness: Empirical and Conceptual Questions. Cambridge, MA: MIT Press.
- Mill, J. 1829. Analysis of the Phenomena of the Human Mind. London.
- Mill, J.S. 1865. An Analysis of Sir William Hamilton's Philosophy. London.
- Moore, G. E. 1922. "The refutation of idealism." In G. E. Moore Philosophical Studies. London : Routledge and Kegan Paul.
- Nagel, T. 1974. "What is it like to be a bat?" Philosophical Review, 83: 435-456.
- Nagel, T. 1979. "Panpsychism." In T. Nagel Mortal Questions. Cambridge: Cambridge University Press.
- Natsoulas, T. 1983. "Concepts of consciousness." Journal of Mind and Behavior, 4: 195-232.
- Nelkin, N. 1989. "Unconscious sensations." Philosophical Psychology, 2: 129-41.
- Nemirow, L. 1990. "Physicalism and the cognitive role of acquaintance." In W. Lycan, ed. Mind and Cognition: A Reader. Oxford: Blackwell.
- Neisser, U. 1965. Cognitive Psychology. Englewood Cliffs: Prentice Hall.
- Nida-Rümelin, M. 1995. "What Mary couldn't know: belief about phenomenal states." In T. Metzinger, ed. Conscious Experience. Paderborn: Ferdinand Schöningh.
- Panksepp, J. 1998. Affective Neuroscience. Oxford: Oxford University Press.
- Papineau, D. 1994. Philosophical Naturalism. Oxford: Blackwell.
- Papineau, D. 1995. "The antipathetic fallacy and the boundaries of consciousness." In T.

Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.

- Papineau, D. 2002. *Thinking about Consciousness*. Oxford: Oxford University Press.
- Peacocke, C. 1983. *Sense and Content*, Oxford: Oxford University Press.
- Pearson, M.P. 1999. *The Archeology of Death and Burial*. College Station, Texas: Texas A&M Press.
- Penfield, W. 1975. *The Mystery of the Mind: a Critical Study of Consciousness and the Human Brain*. Princeton, NJ: Princeton University Press.
- Perry, J. 2001. *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press.
- Penrose, R. 1989. *The Emperor's New Mind: Computers, Minds and the Laws of Physics*. Oxford: Oxford University Press.
- Penrose, R. 1994. *Shadows of the Mind*. Oxford: Oxford University Press.
- Place, U. T. 1956. "Is consciousness a brain process?" *British Journal of Psychology*, 44-50.
- Putnam, H. 1975. "Philosophy and our mental life." In H. Putnam *Mind Language and Reality: Philosophical Papers Vol. 2*. Cambridge: Cambridge University Press.
- Putnam, H. and Oppenheim, P. 1958. "Unity of science as a working hypothesis." In H. Fiegl, G. Maxwell, and M. Scriven eds. *Minnesota Studies in the Philosophy of Science II*. Minneapolis: University of Minnesota Press.
- Rey, G. 1986. "A question about consciousness." In H. Otto and J. Tuedio, eds. *Perspectives on Mind*. Dordrecht: Kluwer.
- Robinson, H. 1982. *Matter and Sense: A Critique of Contemporary Materialism*. Cambridge: Cambridge University Press.
- Robinson, D. 1993. "Epiphenomenalism, laws, and properties." *Philosophical Studies*, 69: 1-34.
- Rosenberg, G. 2004. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. New York: Oxford University Press.
- Rosenthal, D. 1986. "Two concepts of consciousness." *Philosophical Studies*, 49: 329-59.
- Rosenthal, D. 1991. "The independence of consciousness and sensory quality." In E. Villanueva, ed. *Consciousness*. Atascadero, CA: Ridgeview Publishing.
- Rosenthal, D. M. 1993. "Thinking that one thinks." In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Rosenthal, D. 1994. "First person operationalism and mental taxonomy." *Philosophical Topics*, 22/1-2: 319-50.
- Rosenthal, D. M. 1997. "A theory of consciousness." In N. Block, O. Flanagan, and G. Guzeldere, eds. *The Nature of Consciousness*. Cambridge, MA: MIT Press.
- Russell, B. 1927. *The Analysis of Matter*. London: Kegan Paul.
- Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson and Company.
- Sacks, O. 1985. *The Man who Mistook his Wife for a Hat*. New York: Summit.
- Seager, W. 1995. "Consciousness, information, and panpsychism." *Journal of Consciousness Studies*, 2: 272-88.
- Schacter, D. 1989. "On the relation between memory and consciousness: dissociable interactions and consciousness." In H. Roediger and F. Craik eds. *Varieties of Memory and Consciousness*. Hillsdale, NJ: Erlbaum.
- Schneider W. and Shiffrin, R. 1977. "Controlled and automatic processing: detection, search and attention." *Psychological Review*, 84: 1-64.

- Searle, J. R. 1990. "Consciousness, explanatory inversion and cognitive science." *Behavioral and Brain Sciences*, 13: 585-642.
- Searle, J. 1992. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Siewert, C. 1998. *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.
- Shallice, T. 1988. *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.
- Shear, J. 1997. *Explaining Consciousness: The Hard Problem*. Cambridge, MA: MIT Press.
- Shoemaker, S. 1975. "Functionalism and qualia," *Philosophical Studies*, 27: 291-315.
- Shoemaker, S. 1981. "Absent qualia are impossible." *Philosophical Review*, 90: 581-99.
- Shoemaker, S. 1982. "The inverted spectrum." *Journal of Philosophy*, 79: 357-381.
- Shoemaker, S. 1990. "Qualities and qualia: what's in the mind," *Philosophy and Phenomenological Research*, Supplement, 50: 109-131.
- Shoemaker, S. 1998. "Two cheers for representationalism," *Philosophy and Phenomenological Research*.
- Silberstein, M. 1998. "Emergence and the mind-body problem." *Journal of Consciousness Studies*, 5: 464-82.
- Silberstein, M. 2001. "Converging on emergence: consciousness, causation and explanation." *Journal of Consciousness Studies*, 8: 61-98.
- Singer, P. 1975. *Animal Liberation*. New York: Avon Books.
- Singer, W. 1999. "Neuronal synchrony: a versatile code for the definition of relations." *Neuron*, 24: 49-65.
- Skinner, B. F. 1953. *Science and Human Behavior*. New York: MacMillan.
- Smart, J. 1959. "Sensations and brain processes." *Philosophical Review*, 68: 141-56.
- Stapp, H. 1993. *Mind, Matter and Quantum Mechanics*. Berlin: Springer Verlag.
- Stoljar, D. 2001. "Two conceptions of the physical." *Philosophy and Phenomenological Research*, 62: 253-81
- Strawson, G. 1994. *Mental Reality*. Cambridge, Mass: MIT Press, Bradford Books.
- Swinburne, R. 1986. *The Evolution of the Soul*. Oxford: Oxford University Press.
- Titchener, E. 1901. *An Outline of Psychology*. New York: Macmillan.
- Travis, C. 2004. "The silence of the senses." *Mind*, 113: 57-94.
- Triesman, A. and Gelade, G. 1980. "A feature integration theory of attention." *Cognitive Psychology*, 12: 97-136.
- Tye, M. 1995. *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.
- Tye, M. 2000. *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.
- Tye, M. 2003. "Blurry images, double vision and other oddities: new troubles for representationalism?" In A. Jokic and Q. Smith eds., *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.
- Van Gulick, R. 1985. "Physicalism and the subjectivity of the mental." *Philosophical Topics*, 13: 51-70.
- Van Gulick, R. 1992. "Nonreductive materialism and intertheoretical constraint." In A. Beckermann, H. Flohr, J. Kim, eds. *Emergence and Reduction*. Berlin and New York: De Gruyter, 157-179.
- Van Gulick, R. 1993. "Understanding the phenomenal mind: Are we all just armadillos?"

In M. Davies and G. Humphreys, eds., *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.

- Van Gulick, R. 1994. "Dennett, drafts and phenomenal realism." *Philosophical Topics*, 22/1-2: 443-56.
- Van Gulick, R. 1995. "What would count as explaining consciousness?" In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Van Gulick, R. 2000. "Inward and upward: reflection, introspection and self-awareness." *Philosophical Topics*, 28: 275-305.
- Van Gulick, R. 2003. "Maps, gaps and traps." In A. Jokic and Q. Smith eds. *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.
- Van Gulick, R. 2004. "Higher-order global states HOGS: an alternative higher-order model of consciousness." In Gennaro, R. ed. *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamins.
- van Inwagen, P. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.
- Varela, F. and Maturana, H. 1980. *Cognition and Autopoiesis*. Dordrecht: D. Reidel.
- Varela, F. 1995. "Neurophenomenology: A methodological remedy for the hard problem." *Journal of Consciousness Studies*, 3: 330-49.
- Varela, F. and Thomson, E. 2003. "Neural synchronicity and the unity of mind: a neurophenomenological perspective." In Cleermans, A. ed. *The Unity of Consciousness: Binding, Integration, and Dissociation*. Oxford: Oxford University Press
- Velmans, M. 1991. "Is Human information processing conscious?" *Behavioral and Brain Sciences*, 14/4: 651-668
- Velmans, M. 2003. "How could conscious experiences affect brains?" *Journal of Consciousness Studies*, 9: 3-29.
- von Helmholtz, H. 1897/1924. *Treatise on Physiological Optics*. Translated by J. Soothly. New York: Optical Society of America.
- Wilkes, K. V. 1984. "Is consciousness important?" *British Journal for the Philosophy of Science*, 35: 223-43.
- Wilkes, K. V. 1988. "Yishi, duo, us and consciousness." In A. Marcel and E. Bisiach, eds., *Consciousness in Contemporary Science*. Oxford: Oxford University Press.
- Wilkes, K. V. 1995. "Losing consciousness." In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Watson, J. 1924. *Behaviorism*. New York: W. W. Norton.
- Wegner, D. 2002. *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Wittgenstein, L. 1921/1961. *Tractatus Logico-Philosophicus*. Translated by D. Pears and B. McGuinness. London: Routledge and Kegan Paul.
- Wundt, W. 1897. *Outlines of Psychology*. Leipzig: W. Engleman.
- Yablo, S. 1998. "Concepts and consciousness." *Philosophy and Phenomenological Research*, 59: 455-63.